

A Philosophical Justification for the Community of Inquiry Approach to Moral Education

by

Timothy James Sprod

BSc (Tas), PGCE (Wales), MSc (Oxon)

**Submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy**

University of Tasmania (December 1998)

School of Philosophy

Declarations

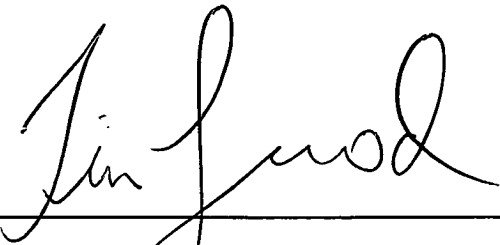
This Thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the Thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgment is made in the text of the Thesis.

Signed



This thesis is made available for loan and limited copying in accordance with the *Copyright Act 1968*.

Signed



Dated

15 APRIL 1999

Abstract

Any society requires that its young be brought up both to live well themselves, and to assist and allow others in the society to live well. In a pluralistic and dynamic society, moral persons must be thoughtful; flexible or steadfast when morality requires it.

Through moral education, schools help their students develop into moral persons. Given that living morally requires a complex mix of moral beliefs, judgements, dispositions and actions, no single pedagogy will be adequate to the task. This thesis considers one pedagogy - the Community of Inquiry - and develops a philosophical underpinning for its use in moral education. This justification will show that, under a particular characterization of morality and moral development, certain capacities essential to moral personhood are best developed through the classroom use of the community of inquiry.

Humans come into the world with genetic predispositions which unfold in transformative interaction with the world: physical and - most importantly - social. Through these, the basis for morality and many other capacities are laid. Chief amongst these is the ability to reason. I argue for a broad based conception of reasonableness, which encompasses five aspects: critical, creative, committed, contextual and embodied. The development of reasoning in social interaction creates persons who can think for themselves through their connection to others. Autonomy is not a cutting off from connection, but an ability to operate competently within intersubjectivity.

The teacher's role in cultivating reasonableness and autonomy can be elaborated through consideration of Habermas' theory of communicative action. I develop a notion of pedagogic action that amalgamates the goal directed strategic educational action of the teacher with the essentially communicative and discursive action through which it must be achieved.

If education is to produce moral persons, we require an account of morality at which to aim. I consider the strengths of Aristotle's virtue ethics, and show that its account of moral development - through habituation - is more reflexive than is commonly thought. Yet, it lacks a detailed account of reflection, and I turn to Habermas' virtue ethics - flawed as a complete meta-ethical system - to provide this, and to inject a dynamic for moral progress.

Finally, I argue that the Community of Inquiry, being a discursive pedagogy, develops the aspects of reasonableness, and builds communicative autonomy in students as they reflectively assess their moral habits and decide in community how to become the sort of person they want to be.

Acknowledgments

The first and most obvious person to acknowledge in a completed Doctor of Philosophy thesis is one's supervisor. In Jay Garfield, I have been blessed. Jay has provided immense intellectual stimulation; detailed, frequent and perceptive feedback on my thoughts and writing; access to many important books in his personal library; and the sort of support and advice without which such an undertaking would be impossible. While I had a reasonably clear idea of where I wanted this thesis to go from the early days, it is a far richer and better argued product for his input. Indeed, sections that I thought I could knock off in a few succinct paragraphs have turned in to pages and pages, but in producing these pages I have learned an immense amount, and strengthened the thesis considerably.

Of course, I have been pig-headed at times, and ignored advice that would arguably have made it even better. But then, it is my thesis, and the responsibility for it rests on my shoulders. For example, I am (I promised Jay that I would make this quite clear) entirely responsible for the decision to avoid gender bias by mixing singular subjects with plural pronouns throughout. I call on such authorities as Fowler and Gowers (1965, 365), who cite such models as Fielding, Goldsmith, Sydney Smith, Bagehot, Bernard Shaw and Ruskin, and the New Oxford Dictionary of English¹ to justify my decision, but Jay still thinks it an abomination. These, however, are minor quibbles. I owe Jay a huge debt for his tireless work on my behalf. He has been a delight to work with and I thank him from the bottom of my heart.

In writing a thesis on moral development, I have had ample reason to reflect on my own upbringing. Hence, I must acknowledge the excellent job done by my parents, Gleewyn and Dan Sprod. In the course of this thesis, I point to the importance of habituation, and I can only say that I was raised into excellent habits. I make clear that habituation is not merely a matter of thoughtless repetitive action, and Mum and Dad understood this well, for I was habituated to be reasonable and reflective throughout my childhood. My parents stood by me then and they have continued to stand by me ever since, offering great support as I threw in a good job and took on the life of a student once again.

That moral development does not ever come to an end is well illustrated by the fact that I also owe an immense debt of gratitude to my wife, Anna Majdanska. Early on in our

¹ In a review in the *Weekend Australian* Review 11, November 14-15, 1998, G. A. Wilkes writes: "However, the usage notes make it clear that 'he' used in reference to a person of unspecified sex 'has become a hallmark of old-fashioned language or sexism in language' and that as 'he or she' can become 'tiresomely long-winded', 'they' as a singular pronoun is to be preferred."

relationship, discussions with Anna started me on the road to the account of reasonableness presented here, by convincing me to look beyond critical thinking. I was a narrow thinker in those days, and continued conversation throughout the years have widened my outlook considerably. Anna has also helped me to see that the link between moral thinking and reflection, and moral action is of vital importance, and that morally good actions depend on an integration of all the aspects of reasonableness and moral knowledge. We can reflectively improve our moral habits at any stage of our lives. I am still working on my interactions with shop assistants!

My debt to Anna is by no means only intellectual and moral. She has been my ideal friend, companion, lover and support. In this last category, she, more than anyone, has borne the brunt of the years as a student, being the mainstay of our family throughout. In thanking Anna for this, I also need to include my children, Liam, Owen and Corin, who have also been a great support to me, and who also cheerfully put up with a father who didn't bring home the bacon. I hope that in future years, they will also look back on their upbringing with gratitude (they could certainly not have wished for a more capable and wonderful mother), but I have learned a great deal from them, too.

The School of Philosophy at the University of Tasmania has been an ideal environment for this work. In particular, I need to thank my fellow postgraduate students and the staff of the School. Working in an intellectually stimulating environment does wonders for one's own work. The stimulation comes from many sources, but I must mention the School's research seminars and the postgraduate work-in-progress seminars. In both these settings, I have not only been introduced to many ideas outside my own interests and been made to think hard, but I have also been frequently surprised and delighted to find that a paper that I thought quite unconnected to my own work has uncovered new directions for me to follow. These formal interchanges, valuable though they have been, pale into insignificance when compared to the informal discussions that occur daily in the School. At the risk of offending others, I must pick out "The Bucket" (our postgraduate room) and especially Cynthia Townley for many long and exceedingly useful conversations. Kicking ideas around in this way is a great method for sharpening one's thinking. The School's Visiting Scholar program has also provided me with personal access to scholars of the calibre of Robert Solomon, Jane Braaten and Timo Airaksinen, all of whom spent time looking at drafts, offering detailed comments and discussing my project with me.

Universities offer many other avenues for intellectual improvement, of course. My involvement in the Centre for Citizenship and Education at the University, including

many discussions with David Hogan, David Owen, Hugo McCann and Mary Fearnley-Sander, has been very valuable to me. Seminars elsewhere in the University, particularly the Removing the Boundaries series, have proved useful.

The University has provided me with two types of opportunities for outreach to the wider world: travel and electronic access. In the former category, I am grateful for support from the Graduate Research Student Support scheme, which enabled me to attend international conferences in New Jersey, USA and Oxford, UK. Attending and, especially, presenting papers at these gatherings enabled me to test my ideas internationally, and my thesis shows considerable benefit from this. Electronic access, especially through email, has certainly widened the ability of students and scholars in relatively isolated locations such as Tasmania to correspond with scholars throughout the world. I have taken advantage of this, picking the brains of many people in many localities, including Finland (Hannu Juuso and Ari Kivelä), Israel (Jen Glaser), the USA (Michael Pritchard and Jill de Villiers), the UK (Nigel Blake), Sydney (Bob Young) and others too numerous to mention. In frequent exchanges on the Philosophy for Children email discussion list (p4c-list), I have benefited from the comments of many others (a partial list can be found in an article I edited from one of these discussions: Sprod 1997a).

Mention of Philosophy for Children brings me to my last acknowledgment. The Philosophy for Children movement has been a part of my life for the past dozen years, and it was obviously the inspiration for this work. In an international community such as this, I have had many discussions about my work, in person - at Philosophy for Children conferences and individually - and electronically, with many people from the movement. I would like to thank them too.

In summary, it can be seen that, although I wrote this thesis, it is in a sense the product of many minds. I have tried to make reference, either individually or as a group, to all those who have contributed, and I am immensely grateful to every one of them. The last mind through which the thesis passed, however, was mine, and I take the responsibility for any shortcomings it has.

Table of Contents

<i>Abstract</i>	iii
<i>Acknowledgments</i>	iv
<i>Table of Contents</i>	vii
<i>Chapter One: Introduction</i>	1
1.1 Background: Approaches to Moral Education	2
1.2 The Purpose of this Study	6
Part I: Reason and Autonomy: Developmental and Pedagogical Foundations	8
<i>Chapter Two: Biological Realism</i>	8
2.1 Negative constraints	11
2.2 Positive nudges	13
2.3 Constraints and Commands	14
2.4 The nature of evolutionary “essences”	15
2.5 From the biological to the social	17
<i>Chapter Three: The Ethical Agent - Reasonable and Autonomous</i>	19
3.1 Reasonableness	21
3.11 Reasonableness and Critical Thinking	22
3.12 The Aspects of Reasonableness	24
3.13 Reasonableness and Moral Education	65
3.2 Autonomy	67
3.21 Kant on Reason and Autonomy	68
3.22 The Development of Reason	70
3.23 Evidence from Experimental Developmental Psychology	74
3.24 A Reconceptualization of Autonomy	81
3.3 Towards a New Understanding of Autonomy	84
<i>Chapter Four: Pedagogical Action and the Development of Autonomy</i>	87
4.1 Habermas’ Classification of Dialogical Action and the Classroom	87
4.11 Critical Theory and Education	87
4.12. The Key Features of Habermas’ Account	88
4.13 Types of Dialogue in Classrooms	99
4.14 A Critique of Habermas from the Classroom Perspective	103
4.15 Pedagogic Action in a Wider Context	115
4.2 Autonomy Again	116
4.21 Dependence throughout Development	119
4.22 Communicative Autonomy	123
4.3 From Foundations to Moral Theory	125

=====

Part II: Moral Theory and Moral Development.....	126
Chapter Five: A Target for Moral Education.....	126
5.1 Moral Development and Meta-ethical Theory	126
5.2 Aristotle's Virtue Ethics as a Multi-dimensional Meta-ethical Theory.....	130
5.21 Moral Development of Children: Instinct and Habituation.....	135
5.22 The Development of Practical Reasoning	140
5.23 Children and Moral Philosophy.....	144
5.3 Summary	152
Chapter Six: Discourse Ethics	155
6.1 Habermas and Discourse Ethics	156
6.2 A Critical Appraisal of Discourse Ethics.....	162
6.21 What is a Consensus?	163
6.22 How Can We Anticipate Normative Correctness?	167
6.23 Rationality	170
6.24 Is (U) a Purely Formal Moral Principle?	185
6.3 Discourse Ethics: Benefits and Drawbacks.....	189
Chapter Seven: A Meta-ethics for an Ethical Education Program.....	191
7.1 Strengths and Weaknesses	191
7.11 What We Can Learn from Discourse Ethics.....	192
7.12 The Weaknesses of Discourse Ethics.....	194
7.2 A Meta-ethics for the Classroom	196
7.21 Inculcation of Respect.....	197
7.22 Dialogue and Development	199
7.23 Progress in Virtue Ethics.....	200
7.24 Justice as a Virtue.....	204
7.25 Virtue in Discourse	220
7.3 From Meta-ethics to the Classroom	222

=====

Part III: Discourse and ethics in the classroom.....	225
Chapter Eight: Ethics, the School and the Community of Inquiry	225
8.1 Moral Education and Schools.....	226
8.11 The Range of School Procedures with Ethical Import.....	227
8.2 The Community of Inquiry.....	230
8.21 Community	231
8.22 Inquiry	232
8.23 The Community of Inquiry as a Setting for Vygotskian Learning.....	234
Chapter Nine: The Community of Inquiry in the Classroom.....	239
9.1 The Development of Reasonableness	244
9.2 Ethical Development in the Community of Inquiry.....	248
9.21 Judgement to Action.....	254
9.22 Bush Lawyers.....	257
9.3 The Place of the Teacher.....	260
9.4 Indoctrination, Ethical Relativism, Moral Indifference and the Path Between Them.....	263
9.41 Indoctrination.....	264
9.42 Ethical Relativism and Moral Indifference	277
9.43 The Path Between.....	281
9.5 Towards Detailed Practice.....	281
Chapter Ten: Practical Implications	283
10.1 The Preliminary Phase of the Community of Inquiry	285
10.2 The Discussion Phase of the Community of Inquiry.....	289
10.21 Following the Inquiry Where it Leads	292
10.22 Reasonableness through and in the Community of Inquiry.....	295
10.23 Moral Correctness and the Community of Ethical Inquiry.....	308
10.24 Communicative Autonomy in the Community of Inquiry	315
10.3 Classroom Communities of Inquiry	317
=====	
Chapter Eleven: Afterword.....	318
11.1 Reasonableness and Autonomy	318
11.2 Ethical Inquiry.....	320
11.3 Lived Lives.....	321
11.4 Conclusion.....	323
Bibliography.....	324

Chapter One: Introduction

In the lives of children between the ages of four and sixteen or eighteen, schools loom large. Almost all children must attend school for about six hours a day, five days a week, for a considerable part of the year. The school is Western civilization's major institution of socialization. Of course, many other institutions have an impact on children's lives, most notably the family and the mass media. Influences from these other institutions can work in harmony with the school, or somewhat at loggerheads.

In the life of a society, the induction of the young into the society also looms large. Particularly in a modern pluralist society, we are faced with twin dangers: the collapse of society if we fail to equip our young to fit in and take on the burdens of maintaining society; or the raising a completely conformist and unthinking new generation unable to adapt in the face of diversity and change. The major challenges facing the present day West spring from rapid changes and increasing globalization, both leading to all of us continually facing what Robert Hughes dubbed "the shock of the new."

Part of the rapidly accelerating rate of change in society is due to the spread of education, particularly in science and technological fields. A more highly educated populace has produced a stunning array of advances across many domains. Standards of living for the ordinary populace in the developed West are far superior to those of the bulk of humanity across all of history. Yet these advances are not an unmixed blessing. Technology and its associates have brought with it dangers that range (for the masses) from nuclear destruction to ecological catastrophe to military armageddon, and (for individuals) from the prolonging of life without hope for the very sick to ruin induced by economic rationalism.

These are well worn observations, and it is just as common to observe that it does little good, or rather great harm, to give children an excellent education in science, mathematics, language and so on, if they do not also gain the moral vision to make not just effective decisions, but wise ones as well. Increasingly, calls are made for schools to do more in moral education and in the closely related field of citizenship education. Yet, opinions differ markedly on what constitutes good moral education, both in terms of what will actually make a difference (the question of effectiveness) and what difference ought to be made (the question of correctness).

This thesis is an exploration of the philosophical underpinnings for one part of the moral educational effort. I will maintain that classroom discussion has a vital role to play in moral education. We shall see that there are strong theoretical and empirical

grounds for the use of discussion, but that the mode of discussion is also important, and the teacher has a critically important part to play in fostering that mode.

1.1 Background: Approaches to Moral Education

Some argue that moral education is not a task for the school at all: it is a task for parents (Beck, Crittenden et al. 1971, 4). Such a position draws a sharp distinction between the school's role of intellectual education, and the home's role of moral training and character building. In times not too far past, the home was assisted in this task by the church. This fact emphasizes the close relationship that was seen between morality and religion. For many, morality was just a matter of heeding God's law.

In the United Kingdom, the state run school was recruited into this enterprise under the 1944 Education Act. Religious Education was, until the Thatcher/Baker Education Reform Act of 1988, the only compulsory subject in state funded schools. In Australia, and even more so in the USA, the situation was different. Religious Education classes in Australia were run by outsiders, not teachers, due to a strong feeling that the state ought not to support a particular religion. In the USA, the strong separation of church and state kept churches even more at arm's length, so that schools and teachers were often urged to be careful to be value-free, and leave moral education to families and churches. Even when schools were to be involved in moral education, it was to be kept carefully neutral (Raths, Harmin et al. 1978).

Despite these different national policies, the upshot was nevertheless similar. Classroom teachers by-and-large felt that it was not their task to be involved in moral education. The task could be safely left to others: parents, churches, religious education teachers (Elias 1989, viii). In the eyes of some, this has led to recent problems. With the increasing breakdown of the family through higher divorce rates, and the decline in church attendance, they argue, the influences of the family and the church on the moral development of children has decreased. Religious Education classes, always somewhat marginalized, have become even more sidelined and ignored, they say, and hence standards of public morality have been falling (Straughan 1988, 1-2 summarizes this view). It is, of course, very difficult to prove anything in this highly complex and contested area: any claim that standards are falling rests on establishing first exactly which standards are right, and why (Straughan 1988, 4-6).

Indeed, a case can be mounted for the opposite view: that moral standards have been rising. Considering practices that were thought uncontentious in the past - taking children away from Aboriginal families; turning a blind eye to domestic violence and child abuse; treating women, the disabled and many others as second class citizens; acceptance of drunken driving; the White Australia policy and so on - others argue that

our society is more caring and more socially just than it was in the past. Social justice, environmental sustainability and democratic values have been so much to the forefront of public moral thinking in the last decade that they are labelled “core values” in the Studies of Society and the Environment learning area National Statement for Australian schools (AEC 1994).

Whether it is true or not that moral standards have been falling or rising, they have certainly been changing, and the moral certainties of the past have come increasingly under question. When combined with increasing globalization and population mobility, this means that children in the West are growing up in more morally complex times.

The linking of moral education to religious education, and more broadly, morality to religion, may well have led many to conflate the two. As religious belief waned in the last few decades, those who took it that morality is equivalent to God’s law would have concluded that, if they no longer believed in God, then there was no real need to believe in a moral code either. In the *Euthyphro*, Socrates recognized the possibility that belief in the gods is not necessary to morality, for any gods worthy of worship will decree what is antecedently moral. It seems wise, in a time of decreasing religious commitment, not to tie moral education to religious education.

I don’t wish to enter into the debate as to whether the West is in moral decline or revival, which is, as I have indicated, highly contentious. Wherever the truth lies, it is nevertheless true that schools have been widely urged to reconsider their role in moral education (Elias 1989; Havel 1995). The Common and Agreed National Goals for Schooling in Australia (the “Hobart Declaration”) includes the goal “6. To develop in students... j. a capacity to exercise judgement in matters of morality, ethics and social justice” (cited in Macintyre, Boston et al. 1994, 190). Given that moral conduct (involving moral judgements) is needed in all the various facets of human endeavour, it seems foolish not to address moral issues in all areas of education, rather than to consign moral education to a single subject. But what sort of moral education? And to what end?

To give a detailed answer to these questions would be a large undertaking. Many different moral education programs have been used, many more suggested. Yet, if we look at the philosophical underpinnings to which these programs appeal, then we may be able to use a broader brush.

If we turn our attention away from moral education programs, towards philosophical accounts of morality, it might seem that we have gained little, for there are also many

philosophical positions. Singer's (1993) edited volume *A Companion to Ethics* has 10 chapters on different moral philosophies, ranging from egoism to Kantian ethics to virtue theory. We can make a little progress, however, if we consider what role reason plays in each philosophical account. Using this criterion, it is possible to set up a continuum from accounts, like Kant's, in which morality is the product of pure reason, to accounts, like Ayer's emotivism, where reason plays no part at all in marking out the moral. We can see that moral education programs also fall along the same spectrum, with Kohlberg's moral dilemma based program at the rational end (Power, Higgins et al. 1989), and Values Clarification at the non-rational end (Raths, Harmin et al. 1978).

In Kohlberg's scheme, moral development is a progression towards a Kantian view. As children attain higher stages of moral development, their moral judgements become more rational, and are based upon more rational grounds. At the highest stage (Stage 6), moral judgements are Kantian, based on the dictates of Pure Reason and aimed solely at justice. Moral development just is the increasing rationality of moral judgement.

In the Values Clarification movement, on the other hand, there are no rational bases for our values. We just have our values, and moral development consists in becoming clearer about what they are. We must reach down to the depths of our being and discover what our true values are. Values Clarification is non-judgemental about the values of others for, according to proponents of this methodology, there are no grounds for criticizing them.

Taking moral education programs from either end of this spectrum highlights one of the dilemmas of modern Western societies. In our pluralist democracy, strong beliefs in the need to deliver justice to all co-exist with a commitment to tolerance of the diversity within our society. These two values do not sit easily with each other. Kohlberg's account of moral development deals easily with justice, but at the cost of relegating many aspects that we think of as moral (notably, care for those close to us) to a realm other than morality. Values Clarification deals with tolerance equally easily, but leaves us entirely without resources for enforcing just outcomes against the wishes of others.

Yet we do have moral intuitions that both values are important. It is in trying to remain faithful to both that many of the dilemmas of moral education arise. If tolerance is given too much importance, to the detriment of justice, the charge can be made that society will break down, lacking the sort of cohesiveness that any society

needs to survive. If justice is overemphasized, to the detriment of tolerance, we face the charge of privileging the values of the powerful over those of the minorities.

Neither side is without resources for refuting these charges, of course. The accounts can be made considerably more subtle than this, in both cases. In part, I am caricaturing the positions to make these points. Yet the caricatures are not too distorting, for the underlying philosophical doctrines for each position lead to these difficulties. The Kohlbergian account rests on an objectivist foundation, while Values Clarification rests on a relativist one. Objectivist meta-ethical positions are notorious for having difficulties dealing with diversity, whereas relativist ones cannot provide any moral reasons for asserting the correctness of any particular position outside the group to which the relativism is indexed.

These two positions also have implications for how we ought to treat children at different ages. As we shall see in Chapter Five, Kohlberg's account implies that children do not act for moral reasons until they are well into their teens, if at all. Nevertheless, they do go through four premoral stages, and reasoning with them at the level of a stage one higher than their present stage will assist them to develop to the next stage. Although we must reason with children, according to Kohlberg, we do not use *moral* reasons until they are at least to stage 4, because they will not be capable of assimilating the reasoning. In Values Clarification, on the other hand, there is only a minor role for reason at any age, as it is restricted to working out what one's values are. The source of values is not clear: it may be due to habituation within a culture, or it may be that values are deep personal attributes.

This last observation leads to the acknowledgment of another aspect of moral development: habituation. In Kohlberg's scheme, habituation into a moral culture is merely an early distraction, heavily influenced in any case by the limitations on the child's (pre)moral thinking. As the child progresses through the stages, only those habitual aspects of morality which accord with practical reason will be retained. As I have noted, Values Clarification is ambivalent about the role of habituation but, if it does have a major role, the values habituated in early life may be modified by clarificatory processes, but they are unlikely to be lost. In Chapter Four, I will consider two accounts of moral development - Aristotle's virtue ethics and Habermas' discourse ethics - that give habituation an important role as the truly moral basis for moral development, accepted in the moral actions of the young, but able to be modified or even overthrown under the influence of practical reason at a later time.

Clearly, the philosophical account we have of reason will be influential in determining the nature of the meta-ethical view we use to guide moral education, but both in turn

shall also have an impact on the methods we use to achieve the end of producing moral persons. This study aims to clarify the philosophical positions that lend credence to a particular practical method for moral education in schools: the community of ethical inquiry.

1.2 The Purpose of this Study

I will develop a philosophical underpinning for the use of a specific type of discussion in moral education. In order to do this, I will examine a number of philosophical issues that impinge on moral development. But it should be noticed that I do not merely develop a philosophical position and then apply it to education (particularly moral education) and childhood. Rather, as we shall see throughout this thesis, my position will be that, although philosophical positions can certainly guide us as teachers of the young, consideration of the young can also guide us as philosophers. Many of the philosophical positions to be surveyed are flawed because they largely consider persons as already formed, and neglect (to a greater or lesser extent) what is to be a becoming-person. And this is precisely what a child is: a becoming-person. It is, of course, what adults are too, although many adults are further along the journey than many children. Karmiloff-Smith (1992, 26) makes the claim that

a developmental perspective is essential to the analysis of human cognition, because understanding the built-in architecture of the human mind, the constraints on learning, and *how knowledge changes progressively over time* can provide subtle clues to its final representational format in the adult mind (*italics in original*).

Exactly the same sort of developmental, dynamic perspective is essential in philosophy, if we are to reach philosophically sound conclusions.

Part I of the thesis is concerned with the developmental and pedagogical foundations on which to base moral education. In Chapter Two, I consider the genetic or biological substrate available for moral development, and begin to explore the complexity of biological and social factors in interaction. In Chapter Three, the focus of attention is on reasonableness, the name I give to a richer and thicker conception of rationality which, as we have seen, is an important enabling capacity for ethical development. After considering the Kantian origin of the concept of autonomy in §3.2, I begin the task of reconceptualizing autonomy in the light of my discussion of reasonableness. Chapter Four will consider the work of Habermas on communicative action, and use the insights of his account to study how it is that teachers can assist their students to develop both reasonableness and autonomy. This will enable us to revisit autonomy, and characterize it in a way that avoids the common charge that it ignores interpersonal relationships and interdependency.

In Part II, I turn to a consideration of the target that we have in moral education. In Chapter Five, I consider Aristotle's virtue ethics and map out its strengths, before identifying a weakness: the lack of a detailed account of the place of discourse in moral development. For this, I explore Habermas' discourse ethics in Chapter Six, identifying the strengths it has, and the reasons why it cannot serve as a meta-ethical theory on its own. In Chapter Seven I map out a meta-ethical account that draws on Aristotle and Habermas, and can serve to provide guidance for a program of moral education.

Drawing upon the conclusions of Parts I and II, I address the implications for the classroom in Part III. Chapter Eight looks at the general role of the school in moral development, and introduces the community of inquiry. In Chapter Nine, I look in particular at the role of the community of ethical inquiry, showing why this methodology can play a vital role in moral development through its encouragement and strengthening of reflection. Chapter Ten addresses the practical implications in the classroom. Finally, Chapter Eleven sums up the arguments advanced in the thesis as to why discussion in a community of inquiry ought to form a central part of any program of moral education.

Part I: Reason and Autonomy: Developmental and Pedagogical Foundations

Chapter Two: Biological Realism

In writing a thesis concerned with the moral education of human beings, it is as well to start by considering just what a human being is. This is, needless to say, an exceedingly complex question, and I shall by no means attempt a full answer to it, though parts of an answer will appear at various junctures throughout the thesis. Nevertheless, I wish to start by emphasizing one important point: human beings are living animals, and thus have a particular biological instantiation. This seemingly obvious point has nevertheless been ignored (or at least, given scant attention) by some of history's most famous philosophers, and the outcome has been some philosophically misleading conclusions (see, especially, §3.2 on Kant). Consequently, I shall start my thesis with a consideration of the biological reality of humanity.

Baier (1985, 5), in a discussion of human capabilities, makes the following comment:

The inheritance includes 'reason,' a product of animal intelligence plus culture-facilitated self-consciousness, and it includes other such joint mental products of equal importance.

By the phrase "animal intelligence," Baier means to indicate that inheritance of reason we have from our evolutionary history. Such an inheritance places limits on the possibilities of being human. We are not infinitely malleable beings, able to be constructed in any way at all by placing us in a suitable social environment. We bring into the world a body (which includes a brain) that has been shaped by evolutionary pressures over at least 3 billion years. Cosmides and Tooby (1994) point out that reason, at least for the higher primates including the early hominids, evolved in an environment where social ties were of vital importance, and that:

the adaptive problems posed by social life loom large. Most of these are characterized by strict evolvability constraints, which could only be satisfied by cognitive programs that are specialized for reasoning about the social world. This suggests that our evolved mental architecture contains a large and intricate 'faculty' of social cognition. (53-4)

Nevertheless, neither Baier, nor Cosmides and Tooby, mean to say that we are determined by what we bring into the world as a result of our evolutionary inheritance. This is clear from the emphasis that Baier puts on "culture-facilitated

self-consciousness.” Like Baier, I believe that many of our capabilities are “joint mental products,” involving biological and social interaction.

Clark (1998, 81) reminds us that our evolutionary inheritance of reason was selected for under conditions that require quick action-oriented solutions to environmental problems in real time. Hence, he claims, we should be wary of considering this inheritance in too unitary a way. “Biological reason,” he later claims (179), “often consists in a rag-bag of ‘quick and dirty’ on-line stratagems.” The extent to which we can believe in the unity of reason will be addressed at several later points in this thesis (see especially §3.126 and §6.2322).

In this section, I wish to explore the notion that, while humans are complex and wonderfully adaptive beings, their adaptability is somewhat constrained by their evolutionary history. Just how constrained, and just what warrant this gives us for basing philosophical arguments on “human nature” is, of course, a very contentious issue. This I shall also explore.

The body, inclusive of the brain and central nervous system, is so constructed that it reacts in certain ways to the world of experience. Some of these reactions are pure reflexes and have very constrained outcomes, such as the walking and rooting reflexes of the neonate, or (even more so) the beating of the heart. Others are largely reflexive - e.g. the fight or flight reaction that floods our body with adrenalin, producing familiar bodily reactions, emotions, feelings and thoughts. Although largely automatic, this reaction can certainly be modified and perhaps even overcome by training. Yet others such as empathy for others are more subtle - they give us more of a nudge than an imperative order - but they are nevertheless there. Reactions of this final type are flexible and plastic, so that through attention, training or social pressures they can be modified to a greater or lesser extent. Nevertheless, they are also subject to constraints built into them by evolution. Flanagan (1991, 41-42) refers to these inbuilt reactions as natural traits, which:

turn up in some recognisable form regardless of cultural context and historical time, and therefore lie closer to our basic biological and cognitive architecture than... the more socially constructed traits... Legitimate contenders... include the six basic emotions of anger, fear, disgust, happiness, sadness and surprise...; the perceptual input system...; the propositional attitudes (but not their contents); biological sex, sexual desire, hunger, thirst, linguistic capacity, and the capacities to be classically and operantly conditioned, to reason, and to remember.... Natural traits of some sort or other constitute the raw material on which all our determinate and socially various traits are in part constructed.

Damon (1996) emphasizes the importance of considering these natural traits when investigating moral development.

Yet biological processes are central to moral growth: There are key natural dispositions that direct early moral reactions and set parameters for later social influence. Any model of moral growth must define the interplay between natural moral dispositions such as empathy and the social influence that transforms these dispositions into stable systems of moral obligation. (200)

Emotional proclivities towards pro-social behaviour appear so early that psychologists now believe them to be inborn. The most unambiguous of such proclivities is a cluster of emotional reactions known as *empathy* and *sympathy*. These interpersonally oriented affective states create in the child a sense of shared responding that links the child psychologically to others. (204-205, original italics)

To argue that such natural traits exist is one thing. To determine what they are is another. Flanagan makes an attempt to identify what he calls “legitimate contenders” in the passage just cited, while Damon adds empathy and sympathy, but two things can be said about this list. Firstly, the case for each item on the list must be an empirical one. Although it is possible to construct such a list sitting in an armchair, the warrant for the place of each rests on grounds that Flanagan identifies as “turning up in some recognisable form regardless of cultural context or historical time.” Consequently, such fields of study as anthropology, history and cross-cultural psychology must be consulted to provide the evidence to substantiate each claim. Secondly (a point that Flanagan makes in his book), disentangling the biological from the socially constructed inputs into natural traits is an exceedingly difficult task. As soon as any biological trait is exhibited (in many cases, even before that), it will be subject to social reaction and shaping. Damon says, “The child’s developing characteristics turn into a kind of ‘second nature’ that is built upon a dynamic interplay between natural events and social influence” (ibid., 208). Thus they are seldom if ever seen in their pure, biological forms. Any attempt to tease out just which components are biological from those that are socially constructed is a delicate reconstructive task.¹

It is well beyond the scope of this thesis to undertake the task of listing, elucidating and defending the natural traits, or of disentangling the biological from the social inputs. It is important to point out, however, that the existence of such constraints on humans must be taken into account in all areas of philosophical inquiry concerning human beings. Just as it would make no sense to invent a sport that required humans to lift twenty tonne weights unassisted, it makes no sense to construct a philosophical position that makes impossible demands on human capacities. Flanagan makes this point, in the context of moral theory, with his Principle of Minimal Psychological Realism: “Make sure when constructing a moral theory or

¹ Despite the widespread usage of “nature” and “nurture” to mark this distinction, I note that there is nothing unnatural about the influence of culture. As Cosmides and Tooby (1994, 42) point out, such dichotomies as evolved/learned, genetic/environmental, biological/social are false dichotomies. In §7.24, I will explore natural traits (or “human universals”) in further detail.

projecting a moral ideal that the character, decision processing, and behaviour prescribed are possible, or perceived to be possible, for creatures like us" (*op. cit.*, 32).

In taking into account the natural traits, we need to be aware that there are two ways in which such traits can be of philosophical interest. Firstly, as Flanagan points out, they impose constraints on us, by making some things impossible. Secondly, they may play a more positive role by imposing certain ways of seeing or doing things upon us, from which we cannot escape. In the next two sections, I shall explore each of these in turn.

2.1 Negative constraints

It seems we must grant that human capabilities are, in some sense, fixed by the evolved human architecture in a way that makes some things impossible to us. We are just not built so as to be able to do them. A philosophical thesis which requires us to be able to do something that we are not, as a matter of fact, able to do must be false. This is the point of Flanagan's Principle of Minimal Psychological Realism.²

Of interest here are those systems of which we can conceive, but which we could not live up to. This is possible in two ways: we might be capable of following the system, but only imperfectly and to a limited degree, because full compliance requires superhuman abilities, or we might be totally incapable of living up to the system in any way, because the system requires that people operate in a way in which they simply cannot operate. A system of the first sort may set us an ideal which can act as some sort of guide to us. A system of the second sort may also set an ideal, but if it is one that is based on a picture of human beings that is seriously at odds with the way we are, then it sets an ideal that not only can't be lived up to, but at which we ought not even to aim, or try to instantiate in even an incomplete way.

The question of ideals will be an important one in this thesis, for many philosophical systems, especially meta-ethical ones, do set up ideals. It will prove to be important to identify the nature of each of the ideals that I discuss. If an ideal sets a standard of perfection, towards which we ought to strive, even while we realize that we are incapable of ever reaching it fully, then we can be happy to be guided by such an ideal. It is quite a different matter if the ideal sets us on the wrong path altogether. But it would be simplistic to set this up as a sharp dichotomy for, as we shall see, it

² To say that some things are impossible for us to do raises the possibility that there are some things are literally inconceivable to us. It may be that there are systems of reasoning or morality that are quite beyond our comprehension. But, if this is true, then it is pointless to speculate about what they might be like, for even such speculation is beyond us.

is quite possible that an ideal should require that we try to approach it up to a certain (quite possibly not very determinate) point, after which further approach becomes counter-productive. I shall illustrate this by reference to rationality.

Reason, or rationality, is an ideal that will be discussed in several places (e.g. §3.1, §3.22, §6.23) in this thesis. It is clear that humans do not reason perfectly at all times. Hence, it might seem a good idea to develop a philosophical account of ideal reason, towards which we can strive. In judging the nature of this idealization, we can consider whether we become better reasoners by more closely approaching it. Consider this situation: two people are walking through a dry sclerophyll forest on a sunny day. Both know that tiger snakes live in such forests, and that they like to warm themselves on flat rocks when it is sunny. Both recognize the type of forest they have entered. The first draws the conclusion from these bits of knowledge that one ought to be wary when crossing flat rocky areas; the other doesn't. Hence the second person is bitten by a tiger snake, while the first, having reasoned better, spots and avoids one. We recognize that the first person reasoned better in this case.

From such considerations, we might set up as an ideal of reason that good reasoners draw all the conclusions that are implied by their beliefs. But, as Cherniak (1986, 12-16) points out, ideal rationality must not require human beings to expend large effort on tasks with minimal survival payoff at the expense of important tasks. Clearly, avoiding being bitten by a poisonous snake has survival value, but it would be absurd for us to draw every conclusion available to us about the animals and plants in the forest, their likely locations at this time of day, where we could find them if we wanted to, what they would be doing, what the effects would be on every other plant or animal, and so on.³ We would not take a step. There are infinitely many conclusions that could be drawn from a human belief set, and we cannot be more rational if we draw them all. We only need to draw those that are relevant and important within the context we are in to be rational. Drawing too many becomes irrational. The ideal rationality of perfect epistemic agents "has seemed a profoundly inapplicable idealization, not just a harmless approximation of actual human rationality" (Cherniak 1986, 134).

The negative constraints inherent in our existence as embodied human beings thus limit the set of possibilities for ideals in many areas. According to Flanagan (1991, Chapter 1), modern findings on human nature show that there is a vast but not limitless possibility space over which human personality can range, and that these limits must also constrain our philosophical conclusions.

³ And I have failed to mention the inferences about what is happening in many other places at this time, on a day such as this, etc., etc.

2.2 Positive nudges

But there is a further point to make. The existence of natural traits may not only negatively constrain what is possible, it may also provide positive input into the social traits that can be constructed. This second point is stronger and more contentious than the first negative one. Rather than just ruling out certain components of human endeavours, it rules them in.

The positive guidance thesis can be asserted in two versions: a stronger and a weaker. The stronger asserts that our biological inheritance makes certain ways of being in the world mandatory - although it does not tell us just how bounded these are, or how deep they lie in our being. The weaker asserts that it is difficult not to be made to see the world in certain ways, but leaves open the possibility that we can, through socially mediated processes, overcome this biological direction. Cosmides and Tooby (1994, 64) can be seen as supporting this stronger version in the following passage:

These circuits inject certain recurrent concepts and motivations into our mental life, and they provide universal frames of meaning that allow us to understand the actions and intentions of others. Beneath the level of surface variability, all humans share certain views and assumptions about the nature of the world and human action by virtue of these universal reasoning circuits.

Karmiloff-Smith (1992) summarizes a considerable body of research that identifies some possible candidates for the stronger claim, a selection of which are: we must see the world as consisting of objects (67); we must assume certain patterns are faces (119); and (in one that has an obvious moral import) we must “attend... differently to the mechanical and the human worlds... understand[ing] others as *subjects*” (121, original italics).

Of course, as Cosmides and Tooby point out (1994, 66), if the strong claim is true it will be difficult, maybe even in some cases impossible, to even recognize that these biases are at work. Even in cases (like those above) where it is suspected, to establish the strong claim will need considerably more empirical research.

Karmiloff-Smith only endorses the weaker claim that we are inherently predisposed to see the world in certain ways, but says that through a process of redescription of our representations of the world, and through making these redescriptions themselves the subject of reflection, we can modify these views (1992, Chapter 7).

We can find support for the weaker claim in Clark’s (1998, 155-157) idea of

partially programmed solutions - that is, cases where the child’s initial program is set up by evolution precisely so as to allow bodily dynamics and local environmental contingencies to help determine the course and the outcome of the

developmental process. Partial programs would thus share the logical character of most genes: they would fall short of constituting a full blue-print of the final product, and would cede many decisions to local environmental conditions and processes.... The idea of a *partial* program is thus the idea of a genuine specification that nonetheless cedes a good deal of work and decision making to other parts of the overall causal matrix. (original italics)

A partial program does not enforce a certain way of being in the world, but it does predispose the child to a range of particular outcomes. Exactly which outcome eventuates is influenced by contextual matters, including in many cases the social context. In the case of outcomes to do with thinking, I shall explore the roles of context and embodiedness further in §3.124 and §3.125 respectively.

The weaker claim is equivalent to claiming that there are default ways of making sense of the world, including the social world, hardwired into us. These predispositions are not rigid, and we can replace the default in the presence of sufficient evidence that we need to do so. As I shall explore further in §3.124, both these positive nudges and the negative constraints identified in §2.1 help to situate us in the world, and hence have moral import.

2.3 Constraints and Commands

It appears that we can assert that there is (most likely) a larger possibility field in such human endeavours as reason and morality than are capable of being instantiated by humans. This is because the way humans are biologically constituted both negatively restricts what humans can do, and positively guides them towards certain ways of being in the world.

To make the case that certain evolutionarily developed traits are an essential part of our make-up does not help us either to identify those natural traits, nor, even if we had a list, to designate, say, which of them push us towards morally praiseworthy action, which push us towards immoral actions, and which have no moral implications at all. Clearly, some of the candidate natural traits put forward by Flanagan are implicated in morality (anger, for example). But is not a straightforward matter to say that a natural trait like anger is clearly an immoral impulse, for anger can be implicated in morally good actions, such as anger at injustice. Indeed, it may be that any of the natural traits may push either way. So, although we can say that the natural traits provide a substrate for morality, they do not do so in any straightforward way. Even such a trait as fellow feeling or sympathy can be a ground for immoral action, as Hume recognises. Sympathy might lead us to favour those who are close to us, to the moral detriment of others, and so it needs to be corrected for bias to produce the steady or general point of view that Hume identifies as moral (Baier 1985, 159-60).

Such a correction needs to be guided by a concept of the moral, and if this is not available as one of the natural traits, as I have been arguing, then it must be a construction of the “culture-facilitated self-consciousness” that forms the second part of the joint production team referred to in the Baier quote at the start of this chapter. Similar arguments apply to any of the human capabilities I have been considering.

In summary, I have been making the point that any discussion of the way humans are situated in the world cannot merely take account of the ways in which society has shaped us. Nor can it appeal solely to a picture of immutable human nature. Both of these need to be taken into account, both singly and for their interactional effects. If Enlightenment accounts have tended to be based on descriptions of the natural Reason, or moral sense or suchlike intrinsic to human beings, then some post-modern accounts (e.g. Lyotard, 1984) seem to have lost sight of the fact that humans come into the world with certain genetically endowed features, that they are situated not just socially and linguistically, but also in their bodies. These bodies have a given structure and a developmental potential; a very plastic one, of course - one that can be influenced greatly by social factors, and can be constructed in many different ways in response to the expectations and beliefs of self and others. Nevertheless, development in certain ways is ruled out, and certain other features of developmental paths are either required, or strongly selected for, by biological factors.

I shall be developing accounts of several human capabilities (reasonableness, emotions, morality) and in each case, the account will draw upon the two aspects identified here: the biological substrate which places certain constraints and guides on the capabilities, and the environmental and social factors that shape their development. In the latter category, my main focus will be on discussion. However, identifying and characterizing the biological inputs is, as I have indicated, a matter for detailed, exhaustive and difficult empirical study. Since I am neither equipped to do that, nor have it as the central concern of my work, I shall often pass over any such account. Rather, I shall try to elucidate the nature of the social influences on these capabilities, particularly those that are mediated through explicit discursive practices.

2.4 The nature of evolutionary “essences”

To claim that certain features have been “bred” into us by our joint evolutionary history may be taken as making an essentialist claim about human beings. Yet precisely what is claimed in a rejection of essentialism is far from clear. Arguments from natural traits have been used to justify many less than salubrious, not to mention false, claims. Crude examples include the natural inferiority of people with

non-white skins and the irrationality of women. Any account of the biological basis for human capabilities that entails such claims must be false. Yet the position that takes the other extreme is equally implausible. It sometimes seems to claim that we merely need to impose a totally different upbringing on a generation of children, and we can make people whatever we want them to be. There is no underlying humanity with which we need to deal.

To tread a middle path on this matter is to take seriously the importance of *both* the biological and the social. As Aristotle (1980) says in relation to the moral, “Neither by nature, then, nor contrary to nature do the virtues arise in us; rather we are adapted by nature to receive them, and are made perfect by habit.” (1103a25). Of course, to advocate a path between the extremes of the mountains and the sea is not to say whether we should be walking in the foothills, on the beach or on the plains between them. Again, Aristotle makes it clear that the discovery of the mean in the virtues is not a matter of the geometer’s rod - the mean can lie closer to one extreme than the other.

In subsequent chapters, I shall deal with discourse - one of the major means for the social formation of human capabilities. Firstly, I will discuss a few features of the biological input to our capabilities that will dispel some of the cruder versions of essentialism - a fixed, narrow and inflexible characterization of human nature - and hence of rejections of essentialism.

The biological bases for our capabilities were shaped by evolution. As such, they must, either now or in the past, have bestowed on us (as a species) some sort of advantage in survival (or at the very least have been associated with some such advantage). Thus, for example, it seems that they must be tied quite closely to our species’ degree of gregariousness, and that a species that had adopted a much more solitary (or more connected) biological lifestyle would have somewhat different biological substrates. As in all things bequeathed to us by evolution, they differ in detail from one of us to another (not everyone’s eyesight is the same, either). So while one person is prone from birth to express anger more readily, another isn’t - and this is not to say that the first ends up being more irascible than the second, for the social modifications of the natural traits can either reinforce or counteract them. For a good summary of the implications of evolution as a basis for human capabilities, see de Sousa (1987, 80-82).

Although biological inheritances underpin all our subsequent competencies, immediately we are born they begin to be overlaid with cultural interpretations and modifications. Indeed, at least some of these biological inheritances have been

“designed” by evolution to facilitate enculturation, for “without universal reasoning instincts, the acquisition of one's 'culture' would be impossible, because one wouldn't be able to infer which representations, out of the infinite universe of possibilities, existed in the minds of the other members of the culture” (Cosmides and Tooby 1994, 72). Neither are they transparent to us. Being at the heart of the way we are situated in the world (see §3.124 for further discussion), it is more likely that they are amongst those features of us which are most taken-for-granted, most hidden from us. And being, in some sense, inbuilt into the way we are, they are deeply inarticulable, at least until we engage in a good degree of self-reflection. This combination of initial and ongoing haziness with the impossibility of separating them out from the social overlays makes it difficult to pin them down clearly.

At least in the case of the three sorts of capabilities with which I am most concerned (reason, emotion, morality), these biological inheritances have irreducibly social aspects. That is, their specific features have been heavily influenced by our evolution as social beings. This is not a Lamarckian point; rather, it is to say that, given the gregarious lifestyle developed by humans, certain interpersonal features will be advantageous to individuals, making them more likely to survive and reproduce. A strictly solitary species would, for example, seem to be totally unable to evolve any reactions that might be said to be moral while, I argue, reason and emotion would be quite different from their human instantiations. Of course, this is linked to the fact that a purely solitary being would also be unable to evolve a number of other important human characteristics, notably language. We need to note here that the capacity to learn a language itself seems to be built into the genetic makeup of humans (Chomsky 1968).

I will argue subsequently that one (but by no means the sole) key element of the social construction for all these capabilities is the internalization of paradigm scenarios, or roles. According to de Sousa (1987, 332), “Like language, these roles are learned in the context of human intercourse and therefore vary significantly from one culture to another. But like language they are learned on the basis of indispensable biological foundations.” It is in this way that we can account for both the core universality of these capabilities over all humans, as well as the social variability.

2.5 From the biological to the social

I have argued in this chapter that, as biological human beings, we enter the world with certain biologically determined traits that both limit how we are capable of being in the world, and positively encourage us to develop in certain ways. None of

this argument serves to fix the way human beings must be in any narrow sense: there are still a large range of potentialities across a vast possibility space for human development. The exact characteristics and capabilities that will be developed in any given individual depend on a myriad of factors, including the broad cultural setting, the family, the specific events experienced, the decisions and judgements that the individual makes and so on. Thus, while there is plenty of room for diversity, there is not infinite room.

In the next two chapters, I will turn the main focus of my attention away from this biological substrate, and explore the philosophical underpinnings for two key characteristics of the ethical agent: reasonableness and autonomy. Nevertheless, the embodiment of humans, and the implications that this carries for philosophical accounts I will be developing, will enter into the argument at various junctures.

In §3.1, I will develop an account of reasonableness that goes beyond the critical and creative aspect commonly attributed to reason or rationality. It is because reason and rationality have been implicitly treated as attributes of decontextualized, disembodied “reasoning machines” that the three further aspects of reasonableness which I identify have been somewhat neglected by many educational theorists. If we take seriously the fact that persons are biological entities, then we cannot ignore the place of the emotions, context and embodiedness in reasoning.

In §3.2, I start by looking at the work of Immanuel Kant, the philosophical source of much of the narrow account of reason and persons, with particular emphasis on his account of autonomy. Autonomous beings, in Kant’s account, largely abstract from their embodiedness and hence must make judgements that contain no hint of commitments that spring from their contextual situation in the social world. Towards the end of that section, I will begin to develop a reconceptualization of autonomy that does recognize the situation of biological and social beings.

Chapter Four shall, through a consideration of the work of Jürgen Habermas, investigate the communicative component of the development of reasonable persons, particularly with reference to the classroom and the position of the teacher. Using this basis, I will revisit the concept of autonomy and further refine it.

Chapter Three: The Ethical Agent - Reasonable and Autonomous

Consider a person acting in the world in a way we consider to be ethically correct. In everyday life, we would ascribe to that person ethical agency. In other words, we assume that the site of the ethical judgement that led to that action is that individual, and hence we ascribe moral praise to them. Lying behind this folk ascription are two key ideas: that persons freely choose to do as they do (i.e. they are autonomous) and that they choose to do so (at least, when they act intentionally) for good reasons (i.e. they are reasonable). The Kohlbergian project, with its paradigm of moral judgement as an individual making a principled choice between the two horns of a moral dilemma, well captures this sense. I shall subsequently argue that, although the Kohlbergian picture captures well the sense of reasoned, autonomous choice, it is not on its own an adequate characterization of ethical agency. The accounts of rationality and autonomy that underpin it are too narrow. In this chapter, I shall map out an expanded rationality, which I will label “reasonableness,” and begin to reconceptualize autonomy.

Ranged against this folk analysis of ethical agency is a view which sees the individual as merely being the implementor of a judgement made elsewhere, for example, by the environment in which the individual is situated (Skinner 1971). Such an analysis might be supported by the often repeated assertion by individuals that they feel they “had no choice” in acting morally (Damon 1996), although we will subsequently see (§5.2) that there are alternative explanations for this feeling, based in the moral character of the individual.

Of course, no individual makes a moral judgement free of any influence from the community in which they live. Indeed, the thrust of this thesis is that living and interacting in a community is an essential element in the development of a moral and ethical individual. But it does not follow that because the individual could not have become a person, an ethical agent, without the community, then the individual is nothing more than that community. Pettit (1993, Part II: Mind and Society) argues persuasively for *individualism* (as opposed to collectivism). Individualism, he says, is the doctrine that individuals do have true intentionality which is not overridden by social regularities, although it is consistent with there being social regularities. This intentionality underlies individual actions, including the making of judgements. A belief in individualism is not inconsistent with *holism*, by which Pettit means the doctrine that the ability to display intentionality depends on thought, which in turn depends on membership of, and development within, a community.

On this account, then, when action is taken, it is taken by the individual. This does not exclude the possibility that some morally evaluable actions of individuals are made non-intentionally (say, from a habit that has been inculcated by the community, and never

consciously considered by the individual) and hence do not involve the individual acting as an intentional moral agent. I take it that not all intentional actions are immediately preceded by judgements, but that judgement must have been implicated in the action at some stage. For example, an action may be the result of a habit that one at some time in the past decided consciously to cultivate. Now, though one performs the action in suitable circumstances without deliberation and judgement, from habit, the action is nevertheless intentional.

Individuals do intentionally make ethical judgements. Sometimes these judgements are followed by actions that flow from the judgements. Ethics requires both a “foundation” in the humanity shaped by evolution, and discursive involvement with others in community. In this section, I intend to concentrate on two necessary conditions of being an ethical agent.

The first attribute can be highlighted by considering the following scenario: a young man is faced with the decision as to whether to register for a conscription ballot. In reaching a judgement about what to do, he engages in many activities. For example: he sits and thinks hard about the issues; he consults a variety of sources, written and oral, to ascertain certain facts; he discusses his options with many people, considering their advice and deciding what influence it will have on his own views. All of these activities involve him in trying to make sense of the various factors which will go together to influence his final judgement: whether or not to register.

If we compare his situation to that of a dog about to be “conscripted,” we can see that the young man not only has a choice, he has the wherewithal to make that choice, in a way that the dog does not. How are we to make sense of the ability of this young man to reach such a decision? What sort of characterization are we to give to this capacity? Some of the differences between the dog and the young man that might account for this can be captured in the following words and phrases: rationality, emotional complexity, imagination, the complexity of representation of the environment and the self, discursive engagement. Do all of these have a part to play? Are there perhaps other factors? How do they fit together? In §3.1, I will develop a notion of *reasonableness* to capture this capacity.

Let’s consider again the contrast of the young man and the dog. Suppose they are both picked up and taken to an Army barracks, to be forcibly conscripted into the Army, under the laws of the country. Suppose that they both find a way to escape, and do so. What would be the attitude of the Army in either case? In the case of the dog, they would most likely either let it go (“We can always get another dog”), or they would recapture it and put it in a better enclosure. Unlike the man, however, they would not consider that the dog bore a moral responsibility for the escape. The young man, however, would find himself on trial. This would reflect the Army’s conviction that his escape was

intentionally decided upon by him as an *individual*, acting on his own behalf and bearing responsibility for that judgement and action.

This sense of being an authentic source of judgement and action which has, in the past, been described using the word *autonomy*, forms the substance of §3.2. The concept of autonomy has increasingly come under attack of late for its extreme individualism and denial of human connectedness; an attack which §3.242 below outlines in further detail. While agreeing that there are problems with the concept and the role it has played, I will nevertheless argue that autonomy does play a crucial role in ethical agency. I shall explore its roots in the works of Kant, before drawing on the conception of reasonableness developed in §3.1 to make a first pass at reconceptualizing autonomy. Chapter Four will continue the task through a consideration of the actions of the teacher in the classroom. But first, let's turn our attention to reasonableness.

3.1 Reasonableness

Humans are, above all, thinking beings. That our enhanced ability to think is what sets us apart from the "lower animals" is a commonplace. Just what this enhanced ability is, however, has been analysed in a variety of ways. Unfortunately, different analyses have not used words or terms in consistent ways. Thus, humans are called the "rational animals," yet just in what rationality consists has been characterised in a myriad of ways. Similarly, "reason," "reasonableness," "critical thinking," "ratiocination" and other terms have been used to capture this capacity. Sometimes two different terms are used to mean roughly the same thing (§6.231 documents Habermas' interchangeable use of "reason" and "rationality"); at others, the same term is used by different authors to mean wildly different things. In the narrowest sense, rationality can be defined as the use of formal logic or, more commonly, informal logic. "Critical thinking" has, by some, been expanded so much as to denote all cognition and much that is affective as well (e.g Siegel 1988, see also § 3.11). "Reason" has been defined as in exclusive opposition to the emotions, or to be that which harnesses the emotions, or to take in the emotions as an integral part: each of these stances is illustrated in §3.1231. Many authors use one or more of these terms without much attempt to characterize them at all.

Yet it does not seem possible to extricate ourselves entirely from this verbal thicket by slashing away the terms altogether. They signify, like all philosophically interesting terms, essentially contested notions. In talking about the aims of education, it is necessary to use them, but it is wise to try to indicate the sense in which they are to be used.

I shall argue that the key guiding intellectual aim of ethical education is to produce (in the words of the title of Pritchard's 1996 book) *Reasonable Children*. It is interesting that Pritchard, despite the title, offers no definition of reasonableness, confining himself to "a rough demarcation."

“Reasonableness” seems a promising term to capture the ability that I am targeting for several reasons. Firstly, it denotes an inherently consensual idea (though I don’t wish to suggest by this that only, or even all, ideas about which a consensus has been reached are reasonable - rather, that reasonableness has at its base the attempt to explore the possibility of consensus). We talk naturally of reasoning *with* others, while we are more likely to be rational *about*, or critical *of*, others. Secondly, it seems to have a less heavy reliance on logic, and create more room for affect, than “rationality” or “critical thinking”: it makes more sense to ask a hard-hearted person to be more reasonable than it does to ask them to be more rational or critical. Thirdly, it seems to create more room for multiple acceptable outcomes. “*The rational solution*” seems more natural, whereas “*a reasonable solution*” comes more easily off the tongue.

“The concept of a ‘reasonable person’ lies at the heart of Philosophy for Children,” say Splitter and Sharp (1995, 6), “and, arguably, of education itself and the ideal of democracy. Reasonableness itself is a rich, multi-layered concept.” They continue by pointing out that reasonableness is linked to, but goes beyond rationality, which they characterize as “all-to-often rigid, exclusively deductive, ahistorical and uncreative.” In addition, they note that reasonableness is social, that it is not just process oriented (requiring sufficient knowledge or content as well), and that it has a “dispositional component... [that] helps to bridge the notorious gap between thought and action” (ibid).

In this section (§3.1) of my thesis, each of these factors will be revisited as I develop an account of the concept of reasonableness, explicating the aspects I see it as having. In §3.2, I will tie this conception of reasonableness to an account of autonomy. But before that, I will make a brief survey at some previous characterizations of reasonableness and its close cousin, critical thinking.

3.11 Reasonableness and Critical Thinking

Partly due to Passmore’s influential 1972 article, the teaching of Critical Thinking become a popularly espoused aim of education over the past decade or two (e.g. Siegel 1988; Facione 1989; McPeck 1990). However, what is actually meant by “critical thinking,” and how wide a variety of thinking types the phrase subsumes, are both highly contested issues. In reaction to attacks on Critical Thinking for being too narrow, linear and logical a conception (e.g. de Bono 1985), the characterization has been widened to Critical and Creative Thinking by some theorists (e.g. Lipman 1991; Boostrom 1992), or even to include Caring Thinking as well (Lipman 1995). My primary aim in this section is to develop a theoretical account of the aspects of thinking and, although occasional reference will be made to the educational implications, I will defer a detailed exploration until Part III.

The adjective “critical” has a perfectly good use, which is weakened by using it to try to cover too many things, and this is what the Critical Thinking movement often seems to

do. For example, in the taxonomy of Ennis (1987), “critical” is made to cover, among other things, sensitivity and dispositions to think. What Critical Thinking theorists are particularly interested in, it seems to me, is thinking that is done, in Stebbing’s (1939) words, “to some purpose”; that is, purposive thinking. Hence, I am excluding from consideration thinking that seems not to have any particular purpose, such as daydreaming, musing, free-associating and so on. I don’t want to make too much of this distinction, though. I realize that such seemingly aimless thinking might serve purposes, such as helping in creative thinking.

The phrase “Critical Thinking” can be ambiguous, in that it can be used to label thinking that is praiseworthy for its quality (i.e. it is being used evaluatively), or thinking that is characterized by certain procedures (i.e. it is being used descriptively). In education, teachers are trying to improve thinking, and to say that one is teaching “critical thinking” is usually taken to mean that one is teaching thinking that is evaluatively good. It needs to be pointed out that critical thinking can be done well or badly, to good ends or bad. In what follows, I do not here mean to imply that all purposive thinking, or any particular aspect of it, is necessarily evaluatively good. My intention is to analyse the aspects of reasonableness, regardless of whether they are well used or not. In other words, I will take as my main aim in the rest of this section a descriptive, rather than evaluative, account of reasonableness.

Of course, since in education we wish to strengthen and improve thinking and reasonableness, this aim shall sometimes enter into the discussion. Often, to attribute reasonableness to a person is to make a positive evaluation of their thinking. Given any particular example of thinking, it can be asked: “Was this thinking good?” If this question is answered in the negative, then (taking the evaluative sense of reasonable) the thinking did not exhibit reasonableness. Of course, the judgement of reasonableness is itself open to critique. Further, the word “good” is in need of some unpacking. How high a standard does it set? Real people who are constrained by limitations on processing and memory capacity, in real contexts where full information about both the situation and the possible consequences is not available, are not capable of ideally optimum performance. The standards need to be set in ways that take these limitations into account (Cherniak, 1986).

In addition to clarifying the standards of performance, the unpacking also needs to make clear the criteria for ascribing a positive evaluation. One of the differences between the connotations of the terms “rationality,” “critical thinking” and “reasonableness” that I mentioned above was that we tend to expect critical thinking to converge on a single rational right understanding, whereas it makes more sense to talk of a range of reasonable interpretations. The criteria for reasonableness will not amount to hard and fast rules for deciding if a conclusion is reasonable. Lipman (1991, 16) makes the point that, on questions that have to do with human conduct, especially ethical matters, we cannot deal

with matters “with the precision characteristic of science. Approximations are needed... We must be content to reach an equitable solution, not necessarily one that is right in all details.” It is possible for reasonable people to agree that a number of options, not all of them compatible, are reasonable. Equally, it is possible for reasonable people to disagree about what would count as reasonableness. Compared to rationality, then, reasonableness is fuzzy. The terms “critical thinking” and “rationality” tend to lead us to expect that, for even humanistic questions, there exists a single right answer if we only had long enough to find it. I will address this question as it arises in the work of Habermas in §6.22.

3.12 The Aspects of Reasonableness

I propose that reasonableness, when looked at from a number of perspectives, reveals five aspects: critical, creative, committed, contextual and embodied. Roughly speaking, they address the calculative, the imaginative, the affective, the situated and the whole body aspects of thinking. An autonomous thinker will be one who is capable of accessing each of the aspects, and of coordinating them, in a manner appropriate to the situation. Because of the large amount of attention that has been lavished on the first two aspects (and not because they are not important), my account of the first two will be relatively brief, and I will concentrate on the latter three, relatively neglected, aspects of thinking.

I want to make it clear that I am not claiming there are five types of thinking. Any specific example of thinking will exhibit each of the aspects, though to greater and lesser degrees. For example, work that we might think is purely creative, like writing a poem, reveals each of the aspects. To illustrate, we see its critical aspect in the judgements involved in choosing one word over another, its committed aspect in the impetus to write it at all, its contextual aspect in the choice of content for the poem or indeed in access to the genre of poetry writing, and its embodied aspect in the metaphoric projection of image schemata (Johnson 1987, Introduction, see §3.125).

Indeed, because the aspects are merely analytic, and not separate thinking types, it is often hard to unentangle them. Embodied thinking involves the physical correlates of emotional states, the emotions we feel are shaped by the context of a situation we are in, the judgement that we are in a particular situation involves critical inferences and so on. The question of the symbiosis between the several aspects will be further addressed in §3.126.

One important point to make about the interrelation of the five aspects is that it is often the case that thinking which is dominated by one aspect can be turned upon thinking heavy with another aspect, or even the same aspect. Thus we can be critical about our critical thinking, or creative about our committed thinking and so on. This reflexivity of thinking is often called *metacognition*, and is important in improving thinking. The important role of metacognition will also be taken up in §3.126.

Thinking, both during the course of and at the end of an inquiry, issues in judgements (Lipman 1991, 65). Judgement also contains an admixture of the critical, creative, committed, contextual and embodied aspects. Judgements that are explicit, conscious and reflective tend to be those that contain a stronger element of the critical aspect of reasonableness. It is probably for this reason that critical thinking has received greater attention than the others. Nevertheless, it must be remembered that judgements with any balance between the aspects can be correct or mistaken.

3.121 The Critical Aspect

Critical thinking has received a great deal of attention. As we saw above, some members of the Critical Thinking movement have characterised critical thinking in a very broad way, to take in thinking that primarily displays what is here called creative or committed thinking as well. However, it seems to me that this is a mistake. One important reason is that the word “critical” carries too many connotations to make it desirable to broaden it like this. For example, it carries a connotation of over-zealous negative fault finding. While giving us pause, such a connotation could be lived with because the word is commonly used in other senses as well. Nevertheless, in talking about critical thinking to those (parents, some teachers, community members, school students) for whom this connotation may be at the forefront of the mind, it would pay to be explicit that this is not what is meant.

More importantly, the etymologically central connections of the word “critical” are to “critique” and “criteria” (Lipman 1991, 116). Indeed, the word “critical” is so tied to “critique” and “criteria” that it seems to broaden it beyond those links, into dispositional or imaginative thinking, is to dilute the term too much. Hence, in this work, I will take it that thinking displays its critical aspect in that it takes what is presently given or proposed and examines it for consistency, the correctness of moves made, the assumptions on which it rests, the implications that may be drawn from it and so on.

The critical aspect to thinking, on this account, does not supply new material to the inquiry, but merely draws out (and makes judgements on) that which is inherent in the material already present. Indeed, to the extent that finding assumptions or implications requires an imaginative act, then these acts involve the supplementation of the critical aspect of thinking by the creative aspect, while the need to actually want to find the assumptions or implications requires input from the committed aspect. This only goes to emphasise that any particular act of thinking can almost always be considered from each of the aspects.

Matthew Lipman, in *Thinking in Education* (1991), advances critical thinking as one of the two components of higher order thinking - the other being creative thinking. Critical thinking he defines as “thinking that (1) facilitates judgement because it (2) relies on

criteria, (3) is self-correcting, and (4) is sensitive to context” (1991, 116).¹ The link of critical thinking to criteria is made clear in this definition, as is that to critique - not just of the content of the thought but also of the thinking itself - through self-correction. Also evident (both here and in Lipman’s definition of creative thinking to be presented later) is the important role of context in thinking. Later, I will make the case that we ought to separate out a contextual aspect of reasonableness.

Because critical thinking has been very widely analysed, I will not spend a lot of time in this thesis developing my own analysis. Different authors come up with different lists of abilities, skills, dispositions, competencies and so on. There are two comments to make about these lists.

Firstly, although these lists differ in detail, there is a great deal of commonality between them. Different conceptual schemata of critical thinking each have their own strengths and weaknesses and a “perfect” scheme is probably impossible to construct. As I have discussed elsewhere, on the basis of an empirical study into the analysis of thinking-in-action in a classroom, the elements of good thinking probably intertwine, rather than form a neat hierarchy (Sprod 1994c), and this makes extricating them very difficult. A list loses the dynamic interplay and contextual deployment of the elements. Nevertheless, listing elements of critical thinking is useful for diagnostic and analytic purposes, and each of the lists can contribute to these tasks.

Secondly, the lists of most authors go beyond what is called the critical aspect to purposive thinking in this work. Take, for example, the list of Ennis (1991, 8-9). Ennis provides a “working definition” of critical thinking: it is “reasonable reflective thinking that is focused on deciding what to believe or do,” and he provides a list of 12 dispositions and 16 abilities that characterise the ideal critical thinker. As an example, the dispositions include “to take into account the whole situation.” Such a disposition (like all dispositions) clearly goes beyond the critical aspect of thinking: it includes the creative aspect in that it requires the exercise of imagination about what might form part of the total situation beyond the immediately obvious; and it requires a desire to look at further evidence - involving the committed aspect. Indeed, the disposition as stated by Ennis is quite impossible. The “total” situation is beyond the capability of any real agent to comprehend (Cherniak, 1986); a rational agent must be able to make decisions about when the amount of information collected and considered is sufficient for reasonable judgement, and this process goes beyond the critical capability. As shall be discussed in further detail below, the emotions play an important part in these saliency decisions (de Sousa 1987).

¹ It is worth noting that Lipman has subsequently added ‘caring thinking’ to the two above, and now describes them as *aspects* and not components of higher order thinking (pers. comm.).

For another example, take Ennis' ability "to ask and answer questions of clarification and/or challenge." While such questions may involve merely identifying the logically puzzling parts of a statement, such as that it involves a contradiction, it is much more often likely to be the case that such questions require an imaginative or creative ability to make connections beyond the given. Equally, there is a desire that lies behind puzzlement, and this lies in the committed aspect to thinking. Each of Ennis' abilities likewise needs more than the mere capacity to carry it out - it is easy to imagine a person who has these abilities but never uses them.

Glaser (1998, Chapter 1) presents a detailed survey of a number of the key players in the Critical Thinking movement. As I did above, she characterizes Ennis' view as "skills plus dispositions," and contrasts this with Paul's "skills plus traits of mind" account and Siegel's "skills plus character" view.² As I have stated, it is not my intention to either survey these accounts in any detail (Glaser's work is well worth reading for this), or to develop a fuller account of the critical aspect to thinking. However, I think it fair to say that each of the extensions beyond skills, and Glaser's own extension to a consideration of identity and the critical thinker, illustrate that there is much more to reasonableness than merely being able to critique.

3.122 The Creative Aspect

One of the clearest indicators of good thinking is that the thinker can go "beyond the information given" (Bruner 1974). Purely critical thinking would be unable to do this, as it is deployed on the information that is either explicitly or implicitly present. Indeed, such a point makes it clear that thinking confined solely to the critical aspect is virtually impossible: all thinking requires input from each of the five aspects. For example, the blending of the critical and creative aspects in all thinking is well caught by Lipman (1991, 68):

There are those for whom good critical thinking is represented by an extreme of analytic precision or logical rationality; others may maintain that really good creative thinking is represented by an extreme of pure intuition or imagination. Nevertheless, if we examine the documents that are considered major products of critical intelligence, we find them shot through with creative judgement. If we examine the works of art that are highly esteemed, we are struck by the amount of sheer knowing and thoughtfulness they contain, not to mention their obvious craft and calculated organization.

Again, there is a considerable literature on creative thinking, though it is not, perhaps, as voluminous as that on critical thinking. Hence, I have no more intention of developing an account of creative thinking than I did in regard to critical thinking. I shall restrict myself to a short survey of a few authors.

Lipman (1991, 193) has a definition for creative thinking - "thinking conducive to judgement, guided by context, self-transcending, and sensitive to criteria" - which echoes

² Glaser herself draws on distinctions made by Siegel (1988, 1-31) in this account.

to some degree that which he gave for critical thinking. Critical thinking, he says, aims at truth, whereas creative thinking aims at meaning. Both, therefore, aim at types of judgement, while the other three features are informative in delineating the differences between the two aspects. Firstly, criteria play a more directive role in critical thinking, which relies on them, as opposed to the creative aspect of thinking, which is merely sensitive to them. Secondly, both the critical and creative aspects of thinking are connected to the context, which (as I shall explicate in §3.124 below) I shall discuss as a distinct aspect of thinking. In Lipman's view, context places constraints on the critical aspect of thinking, but it provides positive impetus to the creative aspect. Finally, critical thinking, as critique, focuses thinking more narrowly towards the acceptable, while creative thinking is expansive and, to echo Bruner, goes beyond by seeking transcendence.

Hence we can see that some of the ways in which thinking displays its creative aspect are that it provides new material for the thinking to grasp, widens the domain of thinking, presents to the mind that which cannot be directly experienced but must be imaginatively constructed and creates links between objects of knowledge previously thought to be separate. By loosening the procedural bounds of critical thinking, it enables more intuitive processes to play a part in thinking. But before continuing to the three aspects of reasonableness that I contend have been somewhat neglected, I wish to highlight two further points concerning the creative aspect to thinking.

The first concerns guiding metaphors for the critical and creative aspects of thinking. Bearing in mind that the two aspects are always blended in any real example of thinking, we can nevertheless see that engaging in thinking in which the critical aspect of thinking is to the fore can be seen as something like writing a set of instructions for a well understood task, whereas thinking in which the creative aspect is predominant is more like telling a story. In the former, the task is mostly to be clear and to get the instructions to fit together properly. The nature of the task compels the nature of the instructions to a high degree. In constructing a narrative, however, there is certainly a need for connectedness - we cannot just put anything next - but there are also many degrees of freedom. The connections need, under some description, to make sense within the greater context, but there are many types of connection. In the instructions, the connections are literal, explicit, but in the narrative they can be drawn from the many literary devices: metaphors, analogies, word associations, rhymes and others.

The second is to note that any discussion of the creative aspect to thinking, especially in a thesis with a focus on ethics, would be incomplete if it did not draw on Johnson's (1993) *Moral Imagination*. Johnson clearly maps out the important place of creative, imaginative thinking in our moral lives and moral theories. He says:

Moral understanding is in large measure imaginatively structured. The primary forms of moral imagination are concepts with prototype structure, semantic frames, conceptual

metaphors, and narratives. To be morally insightful and sensitive thus requires two things: (1) We must have knowledge of the imaginative nature of human conceptual systems and reasoning.... (2) We must cultivate moral imagination by sharpening our powers of discrimination, exercising our capacity for envisioning new possibilities, and imaginatively tracing out the implications of our metaphors, prototypes and narratives. (198)

Further, Johnson makes it clear that one of the key moral sentiments, empathy (about which I will say more in discussing the committed aspect to thinking) is not just a matter of fellow feeling or emotion, but also requires the ability to imaginatively project ourselves into the place of another. It is, he says (*ibid.*, 200) “a blending of feeling, imagination and reason,” and hence (in my terms) involves the committed, creative and critical aspects of reasonableness. Indeed, given that empathy requires a great deal of sensitivity to the context within which our fellows find themselves, as well as an ability to project our whole selves, not merely our thoughts, it will also involve the contextual and embodied aspects (as Johnson - see his 1987 - is well aware).

I shall further draw on Johnson’s account of the place of imagination in moral thinking, and explicate a number of the notions that have been merely indicated here, throughout the remainder of this thesis (see, for example, §3.1253, §6.232 and §7.24).

3.123 The Committed Aspect

A person may be capable of drawing the most obscure implications from a set of information, or of dreaming up the most ingenious of conjectures, yet seldom or never do either. Thus, characterisations of the good thinker usually place a heavy emphasis on the dispositions required, yet little is said about the sources of these dispositions. I assert that it is the committed aspect of thinking that provides the motivation, the engagement of the thinker, and the committed aspect arises in the affective domain. This is not a common view, however, and before expanding on it, I shall first have to tackle the opinion that emotions are the enemy of good thinking. Then, to clear the way for the reinstatement of the emotions into purposive thinking, it will be necessary to examine the nature of the affective domain.

3.1231 Emotional and Rational Thinking

Critical (or rational) thinking is not uncommonly contrasted with emotional thinking. To those who make this distinction, rational thinking is good, because it is clear, cool, calm, considered; emotional thinking is bad because it is muddled, hot-tempered, agitated, rushed. Objective thinking is seen as detached from human desire and emotion, even from individual particularities. Linked to such a dichotomy of reason³ and emotion is a view of “the passions” as something that happens to us, not something we control; of emotions as violent disturbances of our normal smooth running. Such a view has been

³ In this section, “reason” is often used as Hume does, in the sense of unemotional, or anti-emotional, rationality. It can, instead, be used in a much wider sense, to encompass purposive thinking with all its aspects, while “rationality” is reserved for this narrower meaning.

influential in the history of philosophy, versions of it having been advanced by Plato, Descartes, Kant and many others.

Hume (1740/1972, 154-5) points to the ubiquity of this view:

Nothing is more usual in philosophy, and even in common life, than to talk of the combat of passion and reason, to give the preference to reason, and to assert that men are only so far virtuous as they conform themselves to its dictates..... The eternity, invariableness, and divine origin of [reason] have been displayed to the best advantage; the blindness, unconscionable and deceitfulness of the [passions] have been as strongly insisted on.

Hume's discussion of the relation of reason and the passions urges "the fallacy of all this philosophy, [by] endeavour[ing] to prove *first*, that reason alone can never be a motive for any action of the will; and *secondly*, that it can never oppose passion in the direction of the will" (154). In his conclusion to this passage, he famously asserts that "reason is, and ought only to be, the slave of the passions" (156). While this argument clearly reverses the evaluative polarity of the "unruly passions" view, and links the possibility of reason to the passions, it does not seem to deny a clear distinction between the two: they stand in a master-slave relationship. I shall argue for a modified Humean view: that rationality (in the guise of the critical aspect) and emotions (in the guise of the committed aspect) are both essential elements of purposive thinking.

This "divine reason/deceitful passions" view has by no means disappeared since Hume's critique. Many more modern books on clear thinking recommend the separation of the rational and the emotional. For example, Jepson (1952) says of emotional values that "where the facts are in dispute there is no room for them; they cloud and confuse the issue, they effectually beg the question, they disclose prejudice," while Reichenbach (1968) claims that:

the scientific philosopher does not want to belittle the emotions, nor would he like to live without them...but he refuses to muddle emotion and cognition, and likes to breathe the pure air of logical insight and penetration.

A closer analogy for the argument being advanced in this work is a symbiotic relationship, where the two (together with the three other aspects of thinking) are inextricably interlinked. In many cases of symbiosis, an attempt to separate out the two life forms leads to the death of both, their interpenetration is so complex. The unravelling that is attempted in these sections is analytic, but I would assert that any real separation of the aspects of thinking into discrete episodes of purely critical (or whatever) thinking would be just as likely to lead to the death (or, at the very least, the disabling) of each of the aspects of thinking.

This is not to claim that the emotions never interfere with clear thinking, rather that since they are always involved, the proper approach is to acknowledge this and include it in our reflexive approach to improving our thinking. A greater understanding of such an

approach can be seen in the discussion of clear thinking presented by Beardsley (1965, 30):

Clear thinking will be promoted if we can keep the question, "Is such-and-such a religion?" as distinct as possible from the question, "Is such-and-such desirable?" ... It is possible to apply a word according to objective criteria that are independent of the speaker's positive or negative attitudes, as such examples as "mother" and "flag" show. The question is whether common usages are reasonably fixed and clear-cut and whether, when decisions have to be made, the advantages of deciding on one definition rather than another are substantial and plain. Having said this much about the importance of seeking neutral definitions for words having strong emotional associations, we must go on to admit, and indeed emphasize, that emotional and valuational attitudes do indirectly affect decisions to adopt or reject definitions under consideration. It is part of intellectual integrity to acknowledge that this is so and to try to show as openly as possible where it seems to have occurred. (30)

Stebbing (1939) puts the case even more clearly.

I do not in the least wish to suggest that it is undesirable for us to be set on thinking by emotional considerations. On the contrary, nothing else will suffice to make us think to some purpose. Nor do I wish to suggest that the presence of a strong emotion is incompatible with thinking clearly. Certainly, the more strongly we feel the more difficult it is to take account of what is alone relevant. But the difficulty may be overcome, provided that we also desire to reach sound conclusions.... I would say... that it is not emotion that annihilates the capacity to think clearly, but the urge to establish a conclusion in harmony with the emotion and regardless of the evidence. This urge is incompatible with the impartial weighing of evidence which is an essential condition of ascertaining all the relevant facts and deducing the conclusions from these facts alone.

Even Beardsley and Stebbing, however, with their use of phrases such as "objective criteria," "impartial weighing" and "relevant facts" skate over issues such as the place of emotions in influencing what are chosen to count as criteria and facts and the relative weight to be given to conflicting criteria and facts. The unproblematic use of such phrases assumes the possibility of purely rational objectivity, created through abstracting and idealizing an individual, an issue which shall be further explored in §3.2. It is not that emotions can be relegated to being merely the impetus for thinking, or considered as another outside factor that we need to allow for in correcting rational thinking. Emotions are essential to the engagement of thinking. To expand on this point, I shall need to have an adequate account of the emotions themselves.

3.1232 The Affective Domain: Emotions and their Origins

Jaggar (1992) has questioned whether all the elements of the affective domain are sufficiently alike so as to be considered a natural kind. She points out that the wide variety of emotions (instancing the distances between fright and long term dedication; trained aesthetic response and hunger; intense, focused experiences and moods), and the role of social and cultural influences in constituting the emotions, both undermine the suggestion that they can.

Jaggar deliberately spread her net widely: she concedes that some people might not include such feelings as hunger, or moods, in such a list. Nevertheless, she uses the term "emotions" to cover this wide variety. Solomon (1983, 132) also takes a widely inclusive

stance, claiming that “there are three fundamental species of passions”: emotions, moods and desires. For the most part, I shall follow these authors by using the term emotion as a blanket term.

Given the wide variety of entities to be taken in under the blanket, it is not surprising that there are many theories on the origins, objects and classification of emotions. For example, de Sousa (1987) draws on seven modern theories of emotion in constructing his own theory. His account of the objects of emotions lists eleven distinct types of object. Solomon (1983) describes each emotion in terms of thirteen categories, while Haugeland (1985) uses seven categories of types of feeling, each distinguished on six criteria. Such a need for complex descriptions once again underlines the heterogeneity of emotions. I won’t develop and defend a classificatory scheme for the emotions, merely note the complexity of the task. I will, however, outline an account of their origins without, for lack of space, attempting to defend it. This account draws in part on the work of all three authors mentioned above, and will be suggestive for the place of emotion in education.

The first major source of emotions is biological (§2.1 to §2.4). Emotions are clearly linked to physiological events: we have direct experience of these, at least in the case of the violent emotions, such as the “flight or fight” reaction to danger, triggered by the release of adrenalin. The ability to measure physiological changes that occur concurrently with emotional changes led William James to postulate that emotions were nothing but the special perception of the physiological changes. A number of difficulties (notably the inability to locate any mechanism for the perception, and the finding that closely similar physiological arousal can be accompanied by quite different subjective experiences of emotion, even in the same subject) have meant that this project failed. Such phenomena as sexual arousal, triggered in part by the pheromones and by innate reactions to certain visual stimuli, show that in some cases at least the bases for emotional reactions are inherently linked to biology even when the subject of the emotion gives a quite different account of the base. Damasio (1996) refers to such biologically hard-wired reactions as the primary emotions, locating the neural site in the evolutionally old brain. They must have been shaped by evolution and be coded for in our genetic makeup.

In his Chapter 4, de Sousa conducts a detailed investigation of the evolutionary basis of emotions. He makes a number of important points: We cannot assume that, because emotions are biologically based (via evolution), then all humans must have the same basic emotional responses. Evolution creates (and relies on) genetic variation within species. Further, despite the variation in modes of expression, the experiencing of emotion seems to be universal (see also §7.24), and hence it is unlikely that emotions in general were, at the time they evolved, maladaptive (this implies neither that emotions are the optimal solution to the problem they arose to solve, nor that they are necessarily beneficial in particular cases). If emotions are adaptive, they must have had at the time of their

selection (and quite possibly still) a biologically useful function, or be side effects of some other useful function.

The second important source of the emotions is social. While a core repertoire of emotional types seems to be pretty much universal, the experience of emotion - both the specific social circumstances that trigger particular emotions and the modes of expressing emotions - are much more variable and tied to culture. As the anthropologist Lutz (1988, 5)) says:

The claim is made that emotional experience is not precultural but *preeminently* cultural. The prevalent assumption that the emotions are invariant across cultures is replaced here with the question of how one cultural discourse on emotion may be translated into another.... The concepts of emotion can more profitably be viewed as serving complex communicative, moral and cultural purposes rather than simply as labels for internal states whose nature or essence is presumed to be universal.

These are Damasio's secondary emotions, linked inextricably with social and experiential factors. Inextricably, because they are not merely the product of the evolutionally new areas of the brain, but involve the brain sites implicated in primary emotions as well. Damasio (1996, 126) says: "The neurophysiological base of these added strategies is interwoven with that of the instinctual repertoire, and not only modifies its use but extends its reach."

Two particular types of event play a critical part in the establishment of appropriate ways and occasions to experience and express emotions. Firstly, as children grow they observe others around them express emotions in response to the situations in which they find themselves. These scenarios become internalized as paradigms for the children's own emotional reactions. Secondly, and by no means independently, children engage in social interaction on the subject of emotions. Older others give commentaries on emotional reactions the child sees, reinforce the appropriateness or otherwise of the child's own reactions and enter into dialogue with children on the subject of emotions.

Obviously, it is not always the case that these two types of event mesh: it may be a case of "do as I say, not as I do." Nor are the models and pointers received from different people always in harmony. Nevertheless, children grow up with both a rich source of emotional paradigms, and with the knowledge that emotions are negotiable, that they can be modified, suppressed or enhanced in response to reasoned input. They learn that emotional displays have effects on the world and can either help or hinder in the achievement of goals. Much of the cognitive side of emotions comes, it therefore seems, from this social negotiation (by no means solely verbal) of the manner of expression and appropriateness of emotional acts.

3.1233 Emotions, Desires, Values and Commitments

The close link between emotions and desires extends the import of emotions into values and commitments. Each of these three - desires, values and commitments - while not being identical to emotions, contains an essential emotional element.

Solomon (1983, 134) perceptively claims that “the relationship between the emotions and desires is extremely complex, and distinguishing them is not always easy.” Leaving aside “primitive desires” arising from drives such as hunger, he states that “the desires that structure [a person’s] life are... built upon an emotional infrastructure.” Later in his book, however, Solomon seems to take a somewhat different view. He says: “I have argued that every emotion is also an ideology, a set of demands, ‘how the world ought to be.’ It is not only an interpretation of our world but a projection into its future, filled with desires which sometimes become intentions and commitments” (277). By way of contrast, de Sousa (1987, Chapter 8) prefaces his discussion of the time indexing “by first narrowing [his] focus from emotions in general to desire,” explaining that his “excuse... is that desire itself has an emotional aspect, a phenomenological face separable from its functional link to action.” There are several views of the relationship of emotions and desires expressed in these quotes: that desires are built on emotions, that emotions are filled with desires, that desires are a species of emotion, and that all desires have emotional aspects. My claim is that a desire always contains an irreducible element of emotion and that (*pace* de Sousa), it is the emotional attachment that provides the motivational flavour to a desire.⁴

It might be argued that some desires are purely rational, emotionally neutral.⁵ Such reasoning sweeps the emotional aspect under the carpet rather than demonstrating its inapplicability. Emotion is necessary for motivation. As Hume (1739/1962) says, reason alone cannot compel anything. It is the fact that the emotional aspect in cases like this is not violent that enables us to think that it is rationality that provides the impetus.

⁴ One distinction between a desire and an emotion is that a desire must have a propositional content, whereas this is not essential in an emotion. That is, if I desire, I must desire that the state of affairs described by some proposition be met. If I desire a new pen, then I desire that I own a new pen. On the other hand, if I am proud of my pen, then the emotion of pride is tied directly to the pen itself, and not to some state of affairs. Of course, I might be proud that I own the pen, so an emotion can have propositional content, but it need not have. In a desire, I am not indifferent to the propositional element, as I might be in the case of a belief. I can believe that this pen is new without feeling anything for the pen at all: I might just be acknowledging that it meets the criteria of “newness” in pens. If I desire that particular new pen, though, I am emotionally committing myself to the pen to at least some extent. My future emotional state will be better if the propositional part of the desire comes to be. Of course, the desire can be strong or weak, the likelihood of it coming about can be high or low. If the satisfaction of a weak desire is highly unlikely, my emotional attachment can be very small, so that my future emotional state is unaffected by my not obtaining the pen. Nevertheless, if against the odds I did obtain it, there would certainly be an emotional effect.

⁵ An example of such an argument: a dispassionate survey of the options available to me leads me to rationally decide that I ought to desire a new pen. My old pen has just run out of ink, I have an imminent need to write, no-one present has a pen to lend me, I am standing next to a store selling pens, I have the money. These facts, the argument goes, through the pure logic of the situation, mandate that I will desire a new pen. As I argue in the main text, the judgement that a new pen ought to be desired does not, without an emotional element, motivate the actual desiring.

There are certain calm desires and tendencies, which, though they be real passions, produce little emotion in the mind.... When any of these passions are calm, and cause no disorder of the soul, they are very readily taken for the determinations of reason... (ibid., 158)

Indeed, the calm emotions are the longer lasting, the ones that structure our world and our actions within it, that provide us with our general “strength of mind” (ibid., 159). In a “storm” of violent emotion, the strong, reactive, “irrational” emotions can cause us to take actions which, after later reasonable reflection in which the calm emotions infuse our reflection, we recognise as showing poor judgement (being unreasonable). Nevertheless, the person who draws on their strength of mind, their settled emotional dispositions, and resists the temptation to be swept away by the violent emotions is not acting unemotionally, but rather in response to one of a set of competing emotions.

So emotions are an essential element of desires: all desires have emotional aspects. But Solomon claimed that something like the reverse is also true: that all emotions are filled with desires. Emotions do not just concern the way things are presently seen to be, but are also about the way we think things ought to be, or want them to be. We do not just neutrally describe a world unconnected to us; we live in that world and the way of it is intimately connected to the way we are. This is a two way influence: we are the way we are because of the world we live in, but the way that world is for us is a function of the way we are. Our emotions are an essential element of the way we are and of how we are in the world. In this, I am anticipating the fourth aspect of purposive thinking - the contextual aspect - which is closely linked to the committed aspect through the role of emotions in situating us within the world.

A relatively coherent pattern of our desires and preferences is what creates value in our lives.⁶ As Nussbaum (1997) argues, emotions are eudaimonistic: concerned with our flourishing and hence what is important and valuable in our lives. This is not something that the critical aspect to thinking can supply on its own: as Solomon (1983, 126) says, “any attempt to break away from the passions, to become purely ‘objective,’ is already to leave behind all question of value and of meaning.”

The route from emotions through desires and preferences to values leads on to commitment - that which aligns us with particular attitudes, causes and/or actions. Hume is right: it is the settled emotions, that have become like second nature to us through long usage which provide our habitual attitudes and orientation to the world - and they have come to resemble the dictates of reason. Of course, because an attitude is settled does not mean that it is correct or defensible. Yet, though such attitudes and commitments resemble rational beliefs so strongly, it is seldom rational argument that can change them

⁶ Though this might look like a radically individualistic account of value, the pattern of our emotions, desires and preferences itself is heavily influenced by both genetic and social factors, as already discussed. This ensures that human values are not radically individualistic, but it does also explain some individual differences.

- the emotions again need to be involved, often in a more violent form than the settled commitments they are to replace. Emotionally charged experiences make more converts than logical argument.

Yet even logical argument will not be pursued without the emotions. As Stebbing says, nothing but emotion will suffice to commit us to thinking to some purpose. This is the first and primary commitment - to thinking purposively - that must predate any considered judgements and actions in the wider world. When we are committed to our thinking, it is because of the emotions that underpin it. Without any emotion, there would be no commitment. And without any commitment, there would be no reason for thinking rather than not thinking, or for thinking this way rather than that. We must be committed to rationality to imbue our thinking with its critical aspect; we must be committed to imagination or curiosity to imbue it with the creative aspect.

Damasio (1996) provides empirical backing for this claim. He carefully documents cases of specific lesions to the brain that have resulted in the simultaneous, and seemingly intimately connected, strong inhibition of the ability to feel emotions and an inability to plan rationally, or to coordinate and follow larger (particularly social and personal) goals. Interestingly in these cases, it is not that the subjects lose the ability to carry out, on request, such underlying reasoning tasks as generating response options to everyday social situations, considering consequences, conceptualizing the means to achieve social objectives, predicting likely outcomes, or even of achieving normal levels of moral reasoning as assessed by Kohlbergian tests. It is at higher levels, such as being able to place different values on options and outcomes, and in being motivated to carry out the conclusions reached, that they fail. In other words, their thinking can be critical and to some extent creative without being reasonable, precisely because their emotional capacity is damaged, and so their thinking lacks commitment.

Those who denigrate emotional thinking certainly exhibit strong commitment to rationality - yet it is the emotions that generate this commitment. Without commitment, we might allow our thoughts to wash over us, but we would lack the drive to turn them in a particular direction. Purposive thinking therefore requires a committed aspect to mesh with the critical and creative aspects. The committed aspect goes beyond the motivation of thinking itself: it plays a vital role in linking the products of thinking to action, and this is the next issue to be addressed.

3.1234 Judgements and Actions

If the only role of committed thinking was to get the critical and/or creative aspects of thinking started and to keep it going, then the emotions could be banished from the substantive thinking to follow. To come to a rational judgement - say, about what I ought to do in a particular situation - does not entail that I actually carry out that action. Logical conclusions from admitted premises carry no imperative weight; nor do creative solutions

to perplexing problems. There is a need for the will to carry judgements through to action: a will that again arises in the emotion-based committed aspect of thinking. In this subsection, I will explore the connections between emotions, judgements and actions.

Thinking issues in judgements (Lipman 1991), but judgements can vary in many ways. They can be reflective or instantaneous; sound or unsound; primarily critical, creative, contextual or embodied; implemented or ignored. Emotions just are a particular form of judgement, says Nussbaum (1997): they are evaluative judgements about things of importance to us, which acknowledge our neediness and incompleteness before elements of the world that we do not completely control.

Similarly, Solomon (1983; 1988) has argued that emotions are judgements and, further, that all judgements necessarily involve the emotions:

The judgments that constitute emotions must be understood as (1) particular acts of judgement, not just the affirmation of certain propositions, (2) systems of judgments, not single judgments, (3) essentially and not contingently tied to their expression, (4) "dispassionate" only in peculiarly pathological or philosophical circumstances, (5) tied to desires in a non-causal way and (6) sustaining rather than simply initiating, structural rather than disruptive. I will not attempt to establish a sharp distinction between emotional and 'dispassionate' judgments, however, for... I believe that virtually all of our experience is to some degree 'affective', and even our most dispassionate judgments... can be adequately understood only within some larger emotional context. (Solomon 1988, 184)

In these characterizations of the types of judgement that we call emotions, neither Nussbaum nor Solomon is claiming that all emotional judgements are good, sound judgements: on the contrary, any judgement can be unsound. This is the type of emotion seized on by those who wish to contrast "bad emotion" with "good rationality." Unsound emotional judgements tend to be sudden, unreflective and agitated, though examples such as people "working themselves into a self-righteous anger" show that these are not necessary conditions for poor emotional judgements. But then, "dispassionate" critical cognitive judgements need not be sound either: witness the concentration camp guard who reasons that orders from superiors must be obeyed, for such is the analytic nature of orders, superiors and subordinates, but who suppresses the emotions aroused by carrying out the orders. The question of the soundness of emotions as judgements is treated extensively by Solomon (1983, Chapters 12-14), and any education for reasonableness must assist children to be able to assess them.

Several features of Solomon's account above point to the role of emotions in linking judgement and action. He rightly points out in (1) that emotional judgement is more active than a neutral affirmation of "the facts" - we commit ourselves to feeling emotions, as further emphasized by point (3). Nussbaum (1997, 245) also asserts that "judging [is] dynamic, not static." In carrying our judgement from the mental world into the physical, there is an impetus towards further action not inherent in the more purely belief oriented judgements. Point (5) emphasizes the intimate connection of emotions and desires expounded above - and desires are archetypically motives for action. Finally an emotion

(particularly as settled emotion) is typically longer lasting than a thought, and this supports point (6) in its claim that emotions sustain our actions and structure our way of life. Solomon later expands on this point:

The context of an emotion is not just a cognitive context, but an active context in which we are engaged in a world that we care about. To insist on a concept of cognition (or judgment) that is essentially divorced from all such care and concerns is not only to adopt a wholly demeaning conception of emotion; it is also to accept a wholly useless concept of cognition and judgment. (op.cit., 188)

Commitment, based on emotions and desires, motivates our thinking, as we saw in the previous subsection. It is that same impetus that leads to engagement in the world. Emotions provide the attachment - the connection - between judgements and actions. Our emotional reactions are formed (through the construction of paradigm situations) in engagement in the world and they are played out through similar engagement.

Solomon uses the word “care” twice in the quote above to mark this attachment. Lipman (1995; 1997) has added caring thinking to the twin pillars of thinking - critical and creative - advanced in *Thinking in Education* (Lipman 1991). In his characterization, caring thinking is an aspect of higher-order thinking. He describes it as having at least four varieties, which overlap and may not be exhaustive. The first is valuational thinking, which pays attention to what matters in particular perspectives; the second is affective thinking, where emotional reactions play the role of cognitive judgements about appropriateness and justification. Thirdly, there is active thinking, exhibited in looking after (caring for) another, which can be protective or interventionist. One of its manifestations lies in respect for persons: an attitude or a disposition that we bring to our thinking that predisposes us to be careful (care-full) of the others with whom we interact, so as to take their views into consideration, not put them down, support them even if we attack their arguments and so on. Finally we have normative thinking: thinking about what is and what should be, linking the desired with the desirable. This is commonly done through inquiry, emphasizing the reflective component. This normative element, Lipman claims, is always cognitive and inseparable from the other elements of caring thinking. In the next subsection, I will explore the relationship between care, caring thinking and the committed aspect of thinking.

3.1235 Care and Caring Thinking

We have seen that Solomon and Lipman single out care as being important in our engagement in the world. Yet care is a complex notion. In general, to care is to be affected emotionally. A distinction can be drawn between two primary senses of caring, a distinction which is sometimes marked by prepositions: caring *for* and caring *about*.⁷ In the first case, if we care for someone or something, then we take a positive attitude

⁷ Unfortunately, it is not the case that the two senses of care are always marked by the prepositions 'for' and 'about', nor is it the case that when the prepositions are used, they always mark the same distinction. However, I shall use these formulations as a shorthand marker.

towards them or it. We are “for” them. However, to care about something or someone may not imply a positive attitude. If I care about the rising cost of living, this doesn’t imply that I am in favour of the rising cost of living. All it implies is that the issue of the cost of living matters to me; I am not indifferent towards it. We cannot, on the basis of the phrase “care about” alone, decide whether the attitude is for or against. The contrast can be illustrated by the difference in meaning between the following two comments:

1. I don’t care for your attitude towards my dog.
2. I don’t care about your attitude towards my dog.

In the former comment, the speaker is clearly disapproving of the attitude in question; whereas in the latter, the speaker is, as the phrase goes, “beastly careless” - the attitude in question is not a matter of interest, whichever way it goes.

“Care for” is often used to imply more than just a positive orientation, but to include positive actions as well. It is easy to see that the positive orientation is linked to the active “looking after,” for generally when we look after someone, it is because we also have positive feelings for them. Nevertheless, to complicate matters, in institutional settings one who “cares for” another may do it solely for the pay, and not “care about” them at all - or “care about” them negatively. It is curious to note that Lipman (1995), in his own discussion of the “caring for” and “caring about” distinction (he associates them with active valuation and affective valuation respectively), paints care as solely positive and supportive, yet we have clearly seen that care can be negative.

One thing that both the “care for” and the “care about” senses have in common is that they both support the contention that emotions are an essential element connecting judgements to actions. Whether we “care for” or “care about” a critical or creative judgement, the emotional involvement implied by either provides an impetus to action.

Which sense of care is more appropriate to its use in describing an aspect of thinking? If we turn to the emotions that support the movement from judgement to action, we find that they are not all positive towards, and supportive of, others. Strawson (1974) provides a list of emotions that play such a role - resentment, gratitude, forgiveness, love and hurt feelings - while Damon (1988) identifies (partly on the basis of his consideration of empirical studies) empathy, sympathy, admiration, self-esteem, anger, outrage, shame and guilt.⁸

So, to label the affective aspect of thinking “caring thinking” may be seen as trading on this equivocation concerning “care.” Clearly, thinking that is motivated by, say, resentment can be caring thinking in the “caring about” sense. Yet if we talk about caring thinking, others can take us to be drawing on the “caring for” sense and implying that all

⁸ It is worth noting that the context for both Strawson's and Damon's list is *moral* judgement and action.

affective thinking is positive, a position Lipman often seems to take. “Caring thinking” may avoid the connotations of “emotive thinking” or “emotional thinking” - both of which are liable to fall prey to the charge of irrationality that Hume attacked - but “caring thinking” is prone to misunderstanding on the other side. Just as I wish to avoid the implication that all thinking that draws on the critical or creative aspects must be good, I think it wise to avoid the same connotation in the present aspect. Hence the choice of the term “committed thinking,” for this aspect of thinking can commit us to both good and bad thoughts and actions. Committed thinking engages us with the world and others: we care about them.

This is not to say that it is not important to teach children to care for each other in their thinking - of course, it is. But that is to talk to the ways in which the aspects of thinking ought to be used, rather than to analyse what the aspects are.

3.1236 Committed to Thinking

There is another role for the committed aspect to play in our thinking. This is the reflexive one: we need to be committed to the thinking itself. Another way to express this is to say that we need to care for our own thinking - for the correctness of the critical thinking, for the inventiveness of our creative thinking, for the richness of our contextual thinking, for the connectedness of our embodied thinking and for the appropriateness of our selection and integration of the “tools” from the different aspects.

Like other aspects of self-care, the pride we take in our own thinking is intimately tied up with the self-regarding emotions. This is not, however, the whole story. Much of our thinking, especially as we are learning to think better, is done in dialogue, so other-regarding emotions also come into play. A joint commitment to the quality of the group’s thinking is necessary, and it is clear that this is more likely the more emotionally comfortable the group is. We cannot easily separate commitment to the group from commitment to the thinking of the group.

In a sense, we are talking about faith and faithfulness. There is simultaneously a faith in the power of thinking to help us to achieve our aims, and a commitment to the improvement and self-correction of that thinking. These two strands do not always pull in the same direction. In the former we believe that our present thinking is adequate to the task, so that we put our trust in it as it is. In the latter, our faith is in thinking in general, if only we use it properly - a recognition that our present mode of thinking may not be ideal leads us to seek to improve it. Yet there is no ideal way to think - the complexity and richness of thinking ensures that many approaches may be reasonable. So whatever particular style of thinking we use, we must have faith in it, whilst also seeking to improve it. This faith is a type of commitment, and it rests on the values and therefore the emotions that we bring to it. Yet faith goes beyond the emotions, to the way we live

in the world, and hence also draws upon the contextual aspect of thinking as well. It is to this that I now turn.

3.1237 Emotions and the “Frame Problem”

The committed aspect of thinking may be involved in yet another role. One of the major problems in accounts of thinking has been what is called the “frame problem” (e.g. Pylyshyn 1987). In order for thinking to be good, it needs to take account of relevant information, and to exclude irrelevant information. The problem that arises is this: how can thinking know that information not immediately present to it (i.e. not in short term memory) is either needed, or is not needed, if it doesn’t have the access to it to determine that? One of the major roles of emotion, de Sousa (1987, Chapter 7) suggests, is to solve the frame problem. Emotions allow certain information and connections to be highlighted, and others to be neglected, in a process he calls establishing salience. In other words, thinking is committed in one direction over another. Though I think de Sousa’s claim overstates the role for emotions, I shall explore the connection between the committed aspect to thinking and the frame problem. This subsection will lead us naturally into a more promising solution to the frame problem - the contextual aspect to thinking.

When critically studying information, it is possible to draw many, indeed infinitely many, inferences from it. Employing the creative aspect of thinking, it is also possible to intuit many, possibly infinitely many, ways in which the information might be interpreted or linked to other information. Looking for relevant information stored in long-term memory, it is possible to examine for applicability a huge number of facts. As Cherniak (1986) points out in his account of minimal rationality, any thinker attempting to obtain all possible information in such ways would soon run out of resources - time and working memory. They would lapse into paralysis, continuing to think but not getting anywhere, somewhat like a time-sharing computer system that “threshes” when too many demands are placed on it. In the computer case, a program module is needed that allows it to commit itself to certain tasks at the expense of others. According to de Sousa (1987), emotions have evolved to play their part in directing attention, picking out what to concentrate on and what to ignore.

They do this, he says, by “mimicking the informational encapsulation of perception.” Here, he is referring to the way, say, the attention of sight is concentrated on a focal point, with only subsidiary awareness of the periphery, but always able to be switched on the receipt of cues, such as certain types of movements in that peripheral field. Emotions, he claims, tip the balance between conflicting motivational structures. Given the limits on our time, resources and rationality, evolution has provided us with emotions as the mechanism we finite creatures need for terminating such searches. The two “New Biological Hypotheses” advanced by de Sousa (1987) sum this up:

New Biological Hypothesis 1. The function of emotions is to fill gaps left by (mere wanting plus) “pure reason” in the determination of action and belief, by mimicking the encapsulation of perception: it is one of Nature’s ways of dealing with the philosophers’ frame problem.

New Biological Hypothesis 2. Emotions are species of determinate patterns of salience among objects of attention, lines of inquiry, and inferential strategies.

What are we to make of this claim? It is true that emotions do direct salience: we are hypersensitive to clues of danger when we are fearful, for example, or far more ready to notice unusual connections in our knowledge when we are curious. Of course, the emotions do not always appropriately limit and direct our attention. We can be too angry to notice that the person at the focus of our anger wronged us inadvertently, is upset and is trying hard to make it up to us; we can be too grief-stricken to see that Perseus is on the ship with the black sails; or we can be too happy with our lover to recall an important appointment.

Thus, emotions are not a fool-proof solution to the framing problem. Only avoiding framing altogether by a complete search of our knowledge and the environment, together with perfect critical evaluation would even come close to perfection - and we have seen that such an approach is impossible for us. But are emotions the best, or even the only, strategy available to us?

Earlier in this section, I argued that paradigm situations are central to the emotions, drawing there on de Sousa’s account. In §5.1, I shall argue that they are also central to morality. But while paradigm situations are important in both these areas, this is because we frame all our experience in term of situations - as will be argued in §3.124. Emotions clearly have a role to play in the constitution of situations, but it seems to be claiming too much to place the whole burden for framing on them, when other aspects of situations are also important. It is because emotions, seen as paradigms, are species of situations that they can be taken to help solve the frame problem.

3.124 The Contextual Aspect

Thinking takes place in a context. At a surface level, this is obvious enough - a thinker has to think about something. Although, in education, teachers often say that they are concerned with teaching *process* rather than *content*, all that this can mean is that, while they are concerned to ensure that particular processes are learned, they are not concerned about the *specific* content through which they are learned. Almost any old content will do. It is often also assumed that the context of thinking is only superficially important: it can be easily stripped away from the decontextualized thinking skills once they are established, so that they can be used in all sorts of other contexts. The CoRT programs (de Bono 1985), for example, rely on exercises that use contents that vary, seemingly at random, from skill to skill. According to Hennessy (1993), there is little evidence to support the transfer of domain-general skills following the use of CoRT. Indeed, in

general, approaches to the teaching of thinking that, like CoRT, rely heavily on decontextualized skill training have not been particularly successful in creating such general, transferable skills (McMillan 1987; Pascarella and Terenzini 1991, see §10.221).

On the other hand, McPeck (1990) claims that thinking skills must be learned in the content area in which they are to be used. He points to the fact that brilliant thinkers in a certain field seldom seem to be universal geniuses - commonly they are no better than others when they are out of their area of expertise. As we shall see in §3.1242, research on novices and experts, as well as on situated cognition, seems to lend support to such claims. This is an important debate, with many implications for the teaching of thinking and, indeed, for the nature of reasonableness. Though I shall return to it in §3.1245 (see also §10.221), I need first to conduct a closer investigation of the idea of context and the role it plays in purposive thinking.

To concentrate, as the above argument does, only on the content of thinking supplies a very thin context. Thinkers who grapple with the same content area differ in their personal histories, their life situations, their physical surroundings, their hopes for the future, their emotional states, their interests, their reactions to those they are working with, their prior beliefs and concepts, their social status, their gender, their embodiment and in many other ways. This rich context goes far beyond mere content. Theories of the rational agent and of decontextualized thinking ignore all these factors. Like emotions, contextual factors are seen as impediments that need to be overcome, neutralized or bracketed, in order that good thinking can happen. This section will examine the position that, far from being impediments, contextual factors have an essential role to play in the constitution of purposive thinking. The first task, therefore, is to make it clearer what these factors are that go together to make a context.

3.1241 What Constitutes Context?

I have argued that the content of thinking (actual and potential) is insufficient to define its context. In order to explicate what context is, it is worthwhile taking a detour to look at the problems computer scientists have had with creating adequate context for computer-based thinking (or Artificial Intelligence). Proponents of Good Old Fashioned Artificial Intelligence (GOFAI) (Haugeland 1985) assumed that the context of thinking is sufficiently defined by the sum total of the propositions (or “facts”) with which the thinking may be concerned, with the process of thinking being described by a set of domain-general rules for manipulating them (Dreyfus 1992; Clark 1998).

Haugeland (1979) surveys a number of “holisms” - contextual factors that do not seem easy to capture in a list of propositions and rules - that a computer would need to master before it could be said to “understand” natural language. The first - intentional interpretation - involves the computer reacting sensibly to inputs in a way that could lead to an outsider attributing beliefs and intentions to it. A good chess program can achieve

this, within its restricted domain. The moves it makes might lead an outsider to say that it believes that its queen is under attack, and intends to save it by mounting a counter-attack. To a large extent, this holism could be met by a computer that displayed merely the critical aspect to thinking - the application of rules to facts about the positions of the pieces.

Haugeland's second level - common-sense holism - involves the interpretation of ambiguous texts through a common-sense knowledge of related facts. It seems, on the surface, achievable by a machine which has all the possible common-sense connections between facts programmed into it - the connotations, the odd bits of knowledge that come with a word and so on. Computer scientists have tried to achieve it using structured information stores accessed according to inference rules (Dreyfus 1992) and, in small domains, this approach is quite successful. However, the sheer complexity of real world contexts means that the search involved in checking which common-sense tidbits are relevant to extracting meaning in this case must either be arbitrarily terminated or expand to take up inordinate amounts of computing time. People do not engage in large amounts of processing in these situations, yet they are able to disambiguate such examples much more efficiently than computers that arbitrarily terminate. Hence, people must have a different approach.

This problem can itself be subsumed under the next level - the situational. Although it is difficult for a rule-based computer program to discover relevant common-sense connotations and implications of words, the meanings lent to whole phrases and sentences by their situational setting is even more intractable for computers. We all know the feeling of being bemused by what someone has said, having interpreted the situation to be of one sort, only to realise that it is a different type of situation after all. Given our misinterpretation, we have failed to "make the connection" - one that seems obvious once the situation is clear. Once we have a grasp of the situation, we know what is and isn't to be expected within it and how different actors might be interpreting it. The step up in processing demands on a GOF AI computer from the common-sense ambiguity makes these problems harder for computers, but people find it *easier* to carry out a common-sense disambiguation once there is a situation to work in, not harder, even though they have to make sense of the situation as well.

Yet even these situational sites are importantly influenced by the next level again - the existential, where we have knowledge of what it is for both ourselves and others to be actors in the world. Haugeland says:

A single event cannot be embarrassing, shameful, irresponsible or foolish in isolation, but only as an act in the biography of a whole historical individual - a person whose personality it reflects and whose self-image it threatens. Only a being that cares about who it is, as some sort of enduring whole, can care about guilt or folly, self-respect or achievement, life or death. And only such a being can read. (631)

For GOFAI computers to have this, they need to store a huge amount of autobiographical material - but this material is not *them*; it is information, as it were, about some third party, that needs to be accessed anew each time there is a different situation of which to make sense. This creates a representational bottleneck that blocks any possibility of fast, real-time responses (Clark 1998, 21).

Haugeland has been talking about computers and reading: I am interested in people and their place in the world. If we consider the world as text, how do we read it? Like computers, people have memories. But to discover what sort of person I am, I do not need to sift through my memories to find out. I already am that person, and this is present to me immediately in the form of my sense of self. I experience the world *through* my sense of self: I do not neutrally gather information, and later juxtapose it with facts about me. Who I am has an important impact on what I experience, let alone how I go on to make more detached sense of it. That the emotions have a role in setting the existential context is obvious from Haugeland's comment. I do not reconstruct myself anew in relation to every new piece of information - my self is what I bring to that information already; and my self has been constituted by my history, place in society, personality, genetic makeup, present disposition and future aspirations, all in mutual interaction.

Attention to the lessons of studies of child development (Clark 1998, Chapter 2) show that, as we learn about the world, we do not do so through thinking that is carried out alongside, but separately, from our interaction with the world. Neither infants, nor we, gather information neutrally from the world, think about it, then having decided, act on the world. Rather, we engage in "action loops," in which thinking about the world and acting in the world are inextricably linked. Our knowledge is tied closely to our action in the physical world (*ibid.*, 37-39), and is often not available to us if the action we intend to take is different. Indeed, our embodiment in the world is also important here, as I shall explore in §3.125 on the embodied aspect to reasonableness. Clark (*ibid.*, 47) says:

In place of the rational engine cogitating in a realm of detailed inner models, we confront the embodied, embedded agent acting as an equal partner in adaptive responses which draw on the resources of mind, body, and world. (47)

Reference to infants reminds us that we come into the world "always already" situated within it (§2.4). This is due to the fact that each of us is the product of several billion years of evolution, which has equipped us with in-built hardware and software, purposely designed to immediately make some sense of our environment. Cosmides and Tooby claim (1994, 66):

Intuitively, we are all naive realists, experiencing the world as already parsed into objects, relationships, goals, foods, dangers, humans, words, sentences, social groups, motives, artefacts, animals, smiles, glares, relevances and saliences, the known and the obvious.... But to produce this simplified world that we effortlessly experience, a vast sea of computational problems are being silently solved, out of awareness, by a host of functionally integrated circuits. These reasoning instincts are powerful inference engines, whose automatic, non-conscious operation creates our seamless experience of the world.

This is not to say that we come into the world understanding our situation completely: there is still much learning to do, and a good deal of plasticity is built into the system so as to enable us to adapt to the many variant situations in which we might find ourselves (Karmiloff-Smith 1992, 4-6). Yet while we are plastic enough to learn whichever human language is spoken in our environment, we are nevertheless hardwired to learn language (ibid., Chapter 2), and while we can learn the uses and peculiarities of many particular objects, we cannot but see the world as consisting of bounded, cohesive, rigid objects unable to affect each other without touching (ibid., 67).

So when, in §3.1243 below, I talk of being placed in a “totally unfamiliar situation,” we need to be aware that this situation will be only relatively unfamiliar, for we will still comprehend it as consisting of objects, agents, goals and all the rest of the entities mentioned above by Cosmides and Tooby. Brown (1991, 3-4) retells a delightful story from Clifford Geertz’ fieldwork in Bali. Geertz, after making a great deal of the strangeness and difference of Balinese culture, then recalls how he and his wife gained the confidence of the villagers. They were watching an illegal cock-fight when the police arrived. The Geertzes reacted to this exactly as the Balinese villagers did, by scattering, and found themselves elaborately and conspiratorially play-acting an ethnographic interview with a villager when the police came checking. Brown comments that, for all this to happen without any pre-planning or communication, the Geertzes and the villagers, for all their differences, must see the physical and social world in broadly similar ways. Talking of his own fieldwork, Brown (ibid., 151) notes:

Although I started my work with an assumption of cultural difference... it is clear that the rest of the assumptions [I made] are about human psychology: what people in general are prone to do (enjoy praise, dislike blame), and what people under specified conditions will do. Respectively, I had assumed unrestricted and implicational universals. I didn't examine these assumptions carefully for the same reasons that Geertz didn't examine the behaviors he described in order to explain how he had obtained rapport with the Balinese villagers...: because I had no realization of anything that needed explanation. That this said something about the scope and content of human nature - and about the limits of cultural relativity - never crossed my mind, as perhaps it rarely crosses any anthropologist's mind.

This background, consisting of assumptions that are initially provided for us by our biology, and subsequently greatly enhanced by our experience and learning in the world, is what allows us to be “always already” situated in the world.

3.1242 The Situation: Artisans and Apprentices

One of the common analogies for the thinker’s wherewithal is the toolkit. Critical Thinking theorists may talk about the skills, capacities and/or competencies of the thinker as being the tools of thinking, able to be taken out and used effectively as needed. Lipman (1991, 187) points also to the fact that these tools need to be used artfully (as by an artisan) as well - that the appropriate tool must be chosen at the appropriate time, that this choice must fit into a sequence of tool use in which each tool complements the

others, and that the sequence must be guided by criteria appropriate to the task. This raises the question of *how* it is that the thinker knows which tool to use at any specific moment. The same question arises in regard to the knowledge that the thinker needs to draw on. Such coordination is indeed essential, but before considering it further, I'll take a closer look at the use of an individual tool.

Heidegger (1962) and Polanyi (1958) have employed a tool analogy to illustrate this situatedness. They claim that we are situated in the world like an artisan in a workshop. As we use our knowledge of the world, the artisan picks up a tool that lies close at hand and uses it to accomplish a task. The way in which the tool is used is important. If the artisan is experienced and skilled, the tool is not used as a separate entity in the hand according to an explicit set of rules. Polanyi says:

The aim of a skilful performance is achieved by the observance of a set of rules which are not known as such to the person following them. (49, italicised completely in original) ... Rules of art can be useful, but they do not determine the practice of an art; they are maxims, which can serve as a guide to an art only if they can be integrated into the practical knowledge of the art. (50)

Rather, the tool is used as an extension of the body. The user "indwells" in the tool in the same way they do in their own body - it is, in Heidegger's phrase, "ready-to-hand." To become conscious of the tool is to become self-conscious, in the discomfited sense (Polanyi talks of stage-fright), and to be unable to wield the tools in the same, effortlessly skilled way. The artisan becomes the apprentice again, rule-bound and clumsy.⁹ For an apprentice in training, all attention (Polanyi calls this "focal attention") is on the tool itself and the rules of the art. The artisan, by way of contrast, has moved to a new "operational plane" where the task is seen differently - the tool itself receives only "subsidiary attention," while the artisan's focal attention has moved to higher level tasks, such as envisioning the finished product. When instructing an apprentice, many artisans can recall the rules, often with difficulty, but many can't. The great artisan is not always the great teacher - and much of the teaching is done by example and guided practice, rather than explicit instruction. Polanyi says: "An art which cannot be specified in detail cannot be transmitted by prescription, since no prescription for it exists. It can be passed on only by example from master to apprentice" (53).

I referred above to a tool that lies close at hand. A well organized workshop will have the needed tools close at hand, because the situation - the making of furniture perhaps - requires that they be there. In a different situation - a boat-builder's workshop, say - some of the same tools would be there, but some of them would be different. Just so for

⁹ My favorite personal example of this is the occasion in my basketball career when I suddenly thought about which foot I took off from in doing a lay-up. For a while, I was unable to do a lay-up smoothly - I kept tripping over my feet. It was only cured when I went alone to the gym and drilled lay-ups until I had once again forgotten which foot to use, and my feet had remembered. Even now, when coaching beginners, I advise them to drill lay-ups until their "feet think for themselves" - then all they have to think about in a game is whether to go for the lay-up or not. It is hardly surprising that both this example and the example of the tool are physical skills - I will address the embodied aspect of thinking later.

the thinker, according to Heidegger. We recognize the situation that we are in - it has usually been developing around us, and often been shaped by us. The situation is constituted not just by the physical and social surroundings, but also by our goals, our hopes, our concerns, and what we read of the goals of others. Then we lay out the tools (the thoughts, the facts, the emotions, the cognitive moves etc) that are appropriate to the situation - or would it be more accurate to say they come laid out with the situation? This is the question I will address next.

3.1243 The Situation: Whole-part and Part-whole

The puzzle at the conclusion of the previous subsection can be restated: which comes first - the particulars or the holistic situation (or perhaps they come together)? This question of the order can be posed in two different settings: that of the learner and that of the adept.

If we are placed in a totally unfamiliar situation, then we are in the position of the learner who has yet to learn anything. As the situation is unfamiliar, it is clear that we cannot make sense of it in terms of the sort of situation it is. This would seem to imply that it is the particulars in the situation that we must first apprehend. However, many of the particular “tools” we need will not themselves be available to us because, as Dreyfus (1992, 255) says, the “individual features get their significance in terms of an underdetermined anticipation of the whole.” If we don’t comprehend that whole, we don’t see the import of the individual features.

Such a situation is common if we travel or, even more so, if we start to live in another country: it is known as culture shock. When we experience strong doses of culture shock, we are not only at a loss to know what the social situation is that we find ourselves in, we also find it very difficult to “tune into” the cues and components of the situation in order to start to make sense of it. Take the following situation.¹⁰ A newly arrived Asian man approaches the counter in an Australian Government Department. Reading the situation as one of authority-submission, he keeps his eyes politely focused on the counter as he makes a request concerning a form for his elderly and frail mother. The clerk, reading the situation as one of information gathering from a client, reads the averted eyes as indicative of an attempt to mislead. He insists that the mother, disabled as she is, come in and fill in the form herself. The supplicant, reading great disrespect for the elderly, becomes angry just as the clerk, sensing a possible attempt to defraud, also gets angry. Communication totally fails.

As we learn more, these particulars become part of our focal attention, and so are the first things that are seen as we enter the situation. We make sense of the situation haltingly and inefficiently - the clerk who gets some training in cultural awareness is acutely aware of averted eyes, while the acclimatizing Asian man learns to check carefully the intent of

¹⁰ I am indebted to Anna Majdanska for this example.

references to his parents. But as the situation becomes more familiar, the particulars begin to simultaneously take on more meaning and be relegated to subsidiary attention. One day, maybe, the clerk finds himself averting his own eyes when talking to a visiting Asian dignitary, or the Asian man jokes unselfconsciously with workmates about his mother. The unfamiliar object becomes a familiar tool: “Every act of personal assimilation by which we make a thing form an extension of ourselves through our subsidiary awareness of it, is a commitment of ourselves; a manner of disposing ourselves” (Polanyi 1958, 61).

Finally, we become an adept. The situation is now so familiar to us that we “live in it.” At this stage, we can turn our focal attention onto the situation as a whole, with all the skills and tools that we learned present in subsidiary attention - they become, in Polanyi’s words, not the object but the instrument of our attention. Hence, to concentrate more fully on the holistic situation is not to have the particulars disappear, but to sharpen our (subsidiary) awareness of them at the same time. While for the learner, the particulars are the way into the situation, for the adept, they are part and parcel of it.

The above characterization is probably too simplistic, making as it does a sharp divide between a fragmented learner’s world and a holistic adept’s world. In reality, we are placed somewhere along a continuum between the two in many situations. When we feel most at home, when we live in the situation most fully, we are not only most in tune but also the most unaware of the (possibly questionable) assumptions on which we rest. And this is most clearly so when we consider our lives as a whole:

The actual foundations of our... beliefs cannot be asserted at all. When we accept a certain set of pre-suppositions and use them as our interpretative framework, we may be said to dwell in them as we do in our own body. Their uncritical acceptance for the time being consists in a process of assimilation by which we identify ourselves with them. They are not asserted and cannot be asserted, for assertion can only be made from *within* a framework with which we have identified ourselves for the time being: as they are themselves our ultimate framework, they are essentially inarticulable. (Polanyi 1958, 60, italics in original)

Of course, as the phrase “for the time being” shows, this does not mean that we cannot dig down, problematize and investigate parts of this framework from within the framework (see §4.123 for a Habermasian account of this problematization). It is the framework as a whole which is beyond articulation.

An adept of a craft has been through a long apprenticeship, immersed in the workshop. Just as this adept feels “at home” in the workshop, in life, we seldom find ourselves in a totally alien situation. In our “learning” - our engagement in the social world - we have been in many a similar situation. Therefore, the tools that we lay out ready-to-hand are those that we have found useful before, that we have seen other, more competent, artisans use in the past, or that we have used jointly with others, thus making them our own. But they are not laid out once we recognize the situation, because it is through them that we

can make such a recognition, and it is through the situation that the tools have meaning and purchase. They come together as a package. If we change situations, we are translated to a different workshop, and the tools that are ready-to-hand also change.

I have returned to the metaphor of the workshop and talked above about changing from one workshop to another. There is probably need for a word of warning here concerning the workshop analogy. In the real world, we can walk out of one workshop into another - and they can be quite different places. This seldom happens in our metaphorical "workshops": usually, the workshop metamorphoses around us, partly in response to our actions and partly due to other reasons. Of course, it can happen that we are catapulted from one workshop into a quite different one - the example of culture shock discussed above springs to mind. But usually this is not the case.

In relation to the workshop analogy, two broad questions arise. Firstly, if situations are workshops, how small are they? To extend the metaphor, do we have a different workshop for making chairs from the one we use for making stools, or do we only have a different workshop when the objects are described at higher levels - say, furniture and boats. Another way of putting this is to ask, how similar do workshops need to be before we can use the same tools? How different before we need a new set of tools? This is a concern about the distinctness and scope of situations. I will address these questions in §3.1244.

Secondly, we can ask whether there are tools that are useful in all workshops, or a large number of workshops. Do they need minor, or even major, tinkering to make them useful in other workshops? This concern is about transferability across situations and is addressed in the literature of domain-specific vs domain general skills (see, for example, Rogoff 1984; Alexandra and Judy 1988; Perkins and Salomon 1989). While it is dangerous to generalize in this contested area, it does seem that both domain-specific and domain-general skills exist, and that it is possible to generalize the former by widening or merging the "workshops" in which they are used. I will support this claim in §3.1245.

3.1244 Different Workshops: How Big and Distinct are They?

Of course, as indicated above, metaphors like this are bound to break down. If we visited the site of many different workshops - a factory, say - we would most likely find a building with many rooms or, if it were an open plan site, there would certainly be many discrete spaces dedicated to different "work-spaces." Situations are not like that: we seldom walk out a door into a new situation - one situation is much more likely to evolve into another. Similarly, rooms don't overlap (or if they do, the overlap is static), whereas the broader situations we commonly find ourselves in can have features drawn from a number of apparently disparate smaller situations, thrown together in non-systematic ways.

I can be in one situation - say, the well-known restaurant scenario - when I see a good friend laughing and joking with someone I know she despises, or the waiter turns out to share my interest in football, or my companion suddenly turns pale and faints, or... Each of these new situations involves quite different knowledge, skills and abilities from the restaurant situation, but the intersection of the two creates a new and different context. It is not that I have to act, now, according to the restaurant context and, then, according to the other, but that many appropriate responses take the combined context into account. Further, the new, composite situation is subsumed under broader situations - say, how to act in public - and eventually under the broadest situation of all - how to live in the world.

In our everyday life we are, indeed, involved in various "sub-worlds" such as the world of the theater, of business, or of mathematics, but each of these is a "mode" of our shared everyday world. That is, sub-worlds are not related like isolable physical systems to larger systems they *compose*; rather they are local elaborations of a whole which they *presuppose*. (Dreyfus 1992, 14)

Bearing these difficulties in mind, let's return to the issue of the distinctness and scope of situations. The issue raises empirical questions. To what extent do people "wall off" different situations? Do they have one way of coping with a particular situation which is quite different from the way they cope with apparently related situations? How related do situations have to be before they are conceived of as variations on the same situation, rather than as distinct situations? These are difficult questions to answer, as people do not explicitly identify situations as they enter them - rather, they find themselves immersed in them. Like the fish that would be the last to discover water, people must work hard to delineate the situations in which they find themselves. Such knowledge may perhaps come more easily to an outside observer.

One area of empirical research that might help elucidate these questions is research into children's understanding of simple, everyday science phenomena.¹ Children, like adults, act in the world on the basis of their understanding of the situation in which they find themselves. It has been long noted by teachers that many children's apparent understanding of such simple phenomena as light and vision or moving objects do not square with the scientific explanation. For example, children will talk about sight travelling from the eyes to the object the eyes see, or of the need to keep pushing an object in space to keep it moving. A research program has emerged to study such misconceptions (or alternative conceptions). A number of researchers have noted that the context of the way children are asked about these sorts of phenomena can markedly affect the explanations they give.

These explanations seem to arise out of children's mental "paradigm situations." Repeated experience of a particular phenomenon, probably reinforced by the way others talk about it, creates an understanding of the way a small fragment of the world is. The parallels to the *situations* I have been considering is striking. Such research is at an early

stage, however, and there does not seem to be a generally agreed term to cover these fragments. Claxton (1991) calls them “minitheories” and states that:

it is the norm in mental organisation for knowledge and skill to be stored together in purpose-built packages, which have evolved to meet the range of demands which can be expected in a particular type of situation or ‘scenario’. This combination of content, context, purpose and process into flexible bundles of localised competence is the primary format of the mind. Far from being the neat, coherent unitary sort of theory to which science proper aspires, the mind-scape of the child is patchwork and piecemeal. (86)¹¹

Using the term “phenomenological primitives” (or p-prims), diSessa (1988, 50) characterises what appear to be the same constructs as a “visual and inarticulate ability to judge plausibility,” and goes on to say:

though it gives signs of being quite robust, intuitive physics is nothing much like a theory in the way one uses that word theories in... science.... Instead, intuitive physics is a fragmented collection of ideas, loosely connected and reinforced, having none of the commitment or systematicity that one attributes to theories.

The same, or very similar, mental constructs have been identified by Hammer (1996), who also calls them p-prims; Yates et al. (1988), who use the term “prototypes”; and Millar and Kragh (1994), who use the term “context-specific reasoning.” One of the seven “memory elements” White (1988) identifies, namely “episodes,” seems to have similar properties.

All these authors attribute these constructs to repeated common perceptual experiences, and ignore the influence of linguistic or cultural factors, which must also play an important role (Lynch and Jones 1995). These *prototypes* (as I will call them here) are said to be accessed by context (or sometimes analogy), after which they are mentally run to produce predictions.

An example drawn from work on how children make predictions about mirrors, in which I was involved, can illustrate just how restricted these prototypes can be.¹² In investigating the way children and adults made predictions concerning using a mirror (in some cases, to direct a torch beam onto a picture; in others, where to position oneself relative to a mirror in order to see an oblique object), we found that our subjects commonly drew on one of five different prototypes, each of which had several implications:

1. The “seeing yourself in a mirror” prototype.

- 1a Mirrors see what they face.

- 1b Mirrors “point at” what they face.

¹¹ It is interesting to note that Claxton's first sentence claims that “scenarios” are the “norm in mental organisation,” but that by the last, he is confining this claim to the child. I would agree more with the first claim.

¹² I must acknowledge my coworkers on this project - Kevin Collis, Brian Jones, Jane Watson, Sharon Fraser - who contributed greatly to the data gathering and the analysis.

- 1c To see what is in a mirror, you have to be in front of the mirror yourself.
- 1d If you can see something in a mirror, the mirror is “pointing at” it.
- 1e You need to be adjacent to an object to see its reflection.
- 2. The “image as a picture on the mirror” prototype.
 - 2a If you can see the face of the mirror, you can see the image that it has captured.
 - 2b If the mirror’s face can “see” the object, it can capture its image.
 - 2c The object projects its image - if it misses the mirror, then no image appears on the mirror.
- 3. The “mirrors bounce light” prototype.
 - 3a Mirrors can take light around corners.
 - 3b Mirror turns light reliably, but no apparent rule to decide how much - use trial and error to find right angle of mirror.
- 4. The “object and observer symmetry” prototype.
 - 4a Mirror must be “balanced” between source and object (eye and object).
 - 4b More explicit use of angles: e.g. billiard ball analogy.
- 5. The “mirrors redirect active sight” prototype.
 - 5a Sight bounces off mirrors.
 - 5b Line of sight: light and sight, though different, travel the same path; e.g. if you can see object in mirror from the source point, then the source will light up the object.

Not only did all our subjects (child and adult) initially explain their predictions by appeal to one of these prototypes, many of them were happy to jump from one to another depending on the context of the question. As the respondents received feedback from the trials of their predictions, the context within which they were thinking about the task changed accordingly, incorporating the new information. Very few of them tried to look for a coherent prototype or situation into which all of their intuitions fitted: it did not bother them that consecutive statements drawing on different prototypes were contradictory. Even the subject who held an honours degree in science operated (in the

context of our investigation) by jumping from prototype to prototype. It could not be said, however, that these subjects had any difficulty using mirrors in the real world.

I have recounted this research in some detail because it implies that the boundaries of the situations that we find ourselves in can be very narrow. It might be thought that experience with mirrors is a very constrained type of situation - it is certainly many levels of complexity below the restaurant situation - yet it seems that our intuitive understanding of even such a constrained field can be highly fragmented.

However, it is also instructive to consider those who did have a coherent prototype for mirrors. One of our subjects was a 13 year old, who appealed to a “balance” prototype and stuck to it despite changing contexts. His initial statements were quite vague and “non-scientific,” but as he was interviewed, he unpacked these ideas through several versions, each more “scientific” than the former, until his explanation was essentially identical to that taught in schools (though he didn’t use the school jargon). My interpretation is that he had been through a process of making the rules that govern light and mirrors explicit (brought them to focal attention), but that once he has understood them, they have been relegated to subsidiary attention. What remains available immediately to focal attention (in a mirror context) is a slightly fuzzy holistic prototype. Yet this “mirror situation” is able to be unpacked, when needed, into the explicit rules once again - rules that are generally not needed to make sense of mirrors, but which are available when the attention paid to mirrors increases.¹³

Tytler (1994), discussing the mixing by experts of explanations at both the rule-bound, formal science level and the prototype level, comments:

In many cases these multiple explanations represent the intelligent use of conceptions as tools... Such notions [what I am calling prototypes] have a role for the experts as adjuncts, sometimes cues for more complete explanations. The appearance of these notions in everyday language... represent a broader cultural input into the status of these ideas in both novices’ and experts’ minds.... Even for experts, the achievement of powerful, consistent ideas is only a part of the story of the growth in understanding, for more primitive ideas live on in parallel with these. (346)

Education, then, has an important role to play in how individuals frame situations - a number of prototypes can be brought together and “reframed” as a single new prototype, or seen to be “life-world summaries” of formal scientific accounts. The scope and distinctness of a situation is not fixed, but is able to be expanded through reconsideration and reconstruction, including explicit rule learning. A number of situations, previously considered to be discrete, can be subsumed into a new situation of broader scope. Presumably this process is one of the more important in enabling us to learn and to become experts.

¹³ The analysis here bears considerable similarity to Karmiloff-Smith’s (1992, 17-26) account of representational redescription (§3.231), in that we have several explicitness levels of knowledge about the same phenomenon. See also her account of the “child as physicist” in her Chapter 3.

3.1245 Different Workshops: Taking Tools with Us

The second question raised by the idea of different workshops was that of the transferability of tools across workshops. If tools are tied to the workshop within which they are found, then it seems that their use must be learned anew in each new context. The terminology used in discussing this issue varies quite a lot: I shall talk in terms of domain-specific (workshop tied) and domain-general (workshop portable) competencies. After an initial survey of biologically given domain-specific tools, I will assume in this subsection that the domain-specific tools to be discussed are ones that have been constructed through learning within a context.

Many of the domain-specific tools we have are hardwired native modules provided to us by evolution, such as edge-detectors for the visual field (Karmiloff-Smith 1992, 1-4). Some, maybe many, of these modules are “informationally encapsulated,” meaning that they are forever closed to cognitive introspection (only their output is available). Clearly, such domain-specific modules can never be made domain-general. Karmiloff-Smith (*ibid.*, 4-5), however, argues that many domain-specific tools are “modularized as development proceeds,” built on “initial biases or predispositions that channel attention to relevant environmental inputs, which in turn affect subsequent brain development.” This dynamic view of domain-specific modules also allows that, for some at least, the representations of the world they contain can be redescribed and made more domain-general (*ibid.*, 18).

It has been argued that there are only domain-general cognitive competencies. Piaget, the most influential supporter of this view, asserted that when new cognitive competencies are gained (in Piaget’s terms, a new stage is entered), they apply globally, to all contexts (see §3.23 for further discussion of Piaget). His explanation for *decalage*, the failure of a child to use the competencies demonstrated in one domain in others, was that they were still present, but merely obscured by unfamiliar content or contexts. However, this view is probably in terminal decline for, as Karmiloff-Smith (1992, 11) says, “Piaget’s strong anti-nativism and his arguments for across-the-board stages no longer constitute a viable developmental framework.”

Rogoff and Lave (1984, 3) have edited a book of essays on the contextually situated nature of everyday cognition. In the book’s introduction, Rogoff summarizes one of the main implications of the essays it contains: context is an integral part of cognitive competency, not a nuisance variable. She says:

Evidence suggests that our ability to control and orchestrate cognitive skills is not an abstract context-free competence which may be easily transferred across widely diverse problem domains but consists rather of cognitive activity tied specifically to context.... This is not to say that cognitive activities are completely specific to the episode in which they were originally learned or applied. In order to function, people must be able to generalise some aspects of knowledge and skills to new situations.... [There is a need to focus] on determining how generalisation of knowledge and skills occurs.

Rogoff is, of course, correct that it cannot be the case that cognitive competency is uniquely tied to a particular context. If this were the case, then it would be impossible for humans to make any radically new intellectual discoveries, or to teach themselves about new areas. These abilities rely upon domain-general cognitive activity. Nevertheless, her final call for better understanding of how such generalization takes place is important. It is a call that is echoed by Alexandra and Judy (1988, 382). In their review of novice-expert studies; they say, “What seems obvious to us from this review is that the investigation of the interaction between domain specific and strategic [i.e. domain-general] knowledge is very much in its infancy” and give a long list of inadequacies in present research.

Nevertheless, Alexandra and Judy summarize some of the findings that seem to be reasonably well supported. People need to know some domain-specific facts before domain-general competencies can be applied, but then such domain-general competencies do contribute to using and acquiring further contextual facts and skills. As the domain-specific knowledge and competency increases, however, the nature of strategic processing changes, and domain-specific processes become more important. Experts do have specialized knowledge and skills that work better than generalized approaches. The extent to which this happens depends on the nature and structure of the domain and the task: highly structured or rich domains call more on specialized domain-specific competencies and knowledge, whereas it is easier to cope in simple domains by using general competencies. Finally, those who perform best are those who can see how both domain-specific and domain-general competencies are related across different tasks and domains. Such an understanding is built through stepping back and making these competencies themselves the subject of study - that is, through metacognition. Since all the aspects of reasonableness are involved in metacognition, I shall return to this in §3.126.

Studies such as those reviewed by Alexandra and Judy seem to paint a picture of two sorts of tools. The domain-general tools are flexible in that they can be applied to many situations, but the price they pay is that they are not especially well suited to particular tasks. Domain-specific tools, on the other hand, are highly specialized and so well suited to their particular domains, but unable to be used easily in different situations. A further point remains. Several domain-specific tools might differ in detailed applicability, yet more detached reflection might indicate that they are based on, and do the same general task as, a domain-general tool. Lying behind this assertion is the question as to whether context can be widened in this way, so that apparently distinct situations come to be seen as, in important ways, the same situation.

If I push the tool analogy a little further, I can illustrate this point. A spanner set will contain a large number of fixed size spanners, each of which can be used to tighten or undo a nut or bolt of a particular size. A shifting spanner, on the other hand, can be

adjusted to fit all of these but it will more often slip off, because it isn't really possible to get it to stay as rigidly tight on a nut as a fixed spanner - there is more play between the jaws. Nevertheless, the fixed and shifting spanners accomplish the same basic task. If the analogy holds sufficiently - if, for example, cases of inference in science, history, art and mathematics can be considered as being as alike as several different sized bolts - then we can see an argument for the construction of better domain-general tools.

Our experience seems to tell us that the tools of the mind are unlike the spanners in this respect: the mind can remake a tool more easily than a fixed spanner can be remade. If a thinker can come to see that several different tools are doing the much the same task in different domains, then they can construct a generalized tool that has applicability not only across those domains, but also potentially in others yet to be considered. Karmiloff-Smith (1992, 17-26) argues that such representational redescription is the means by which domain-specific competency can be transformed into domain-general competency. Perkins and Salomon (1989) argue that this will not happen unless the conditions for generalizing are in place.

Much of this subsection has tended to conflate the "tools in the workshops" with the critical aspect of thinking. This is partly because much of the work in this area has concentrated on critical thinking to the detriment of the creative, committed and embodied aspects of thinking. The points made, however, can be generalized to these three aspects as well.

Thinking has been taken by many to be a purely mental activity, yet the analogy of the use of tools suggests the impossibility of divorcing the physical and the mental in the act of thinking. It is to the inherently embodied nature of thinking that I now turn my attention.

3.125 The Embodied Aspect

3.1251 Perception and Skilled Motor Activity

Perhaps the best place to start in a discussion of the embodiedness of thinking is with practical examples. In looking at a rock, say a chunk of granite, in order to identify the minerals present, the physical act of seeing the mineral and the mental act of inferring from what is seen to the identity of the mineral are inextricably merged. Without seeing it in the right way, the identification is difficult to impossible - quartz looks too much like some of the feldspars. A knowledgeable person holds and manipulates the rock in a different way from the novice - the rock is held in a particular orientation to the light, and is rocked slightly, so that the light can reflect "rollingly" off the conchoidal fracture of quartz, or "flash" off the cleavage planes of the feldspar. Telling this to the novice is not sufficient for them to be able to do it - they need to also have their attention directed and to practice this holistic ability to see/infer.

Another anecdote on this last point. When I was doing my Geology degree, I studied ore deposits. Identification of ore minerals in polished section was a major part of the practical exam, but in the prac sessions, I found that I could not distinguish one from another, despite closely reading their descriptions. In a panic that I would fail the prac exam, I spent many hours over the microscope, looking at labelled specimens and studying the descriptions. One day, I realized that I could now reliably identify them, but when I tried to explain how to a friend, I could only echo the differences referred to in our notes - differences that now jumped out at me, but which I had been completely unable to see before. There was no change in my explicit knowledge, but I had learned how to see/infer.

The gaining of expertise in the ability to see/infer within specialist areas is a well studied phenomenon. Larkin et al. (1980), for example, cite their own previous studies that found that chess masters & grand masters recognize perhaps 50 000 chess board patterns, the equivalent (they say) of the vocabulary of a college graduate. Clark (1998, 14), citing robotics research, claims that “there is no clear dividing line between perception and cognition, no point at which perceptual inputs are translated into a central code to be shared by various onboard reasoning devices.” Rather, perception, cognition and action in the world are tightly tied together in “action loops” (ibid., 36).

I have used the phrase *see/infer* to emphasize that perception is not a separate business from cognition. The inferences in everyday perception are either hard-wired, or so completely automatized that we do not realize that the inference is being done - we accept our perceptions as given (Karmiloff-Smith 1992, 67). When we learn specialist perception, like that of the geologist or the chess master, the inferences become, for a while, explicit and difficult. Soon, however, they too are automated and we forget that we are doing them. This is one of the difficulties of the teacher - remembering again what it was like not to be able to see. There are two parts to this special vision - the first is learning what to direct your attention towards and is reasonably articulable, but the other is just “being able to see it” and is much harder to make explicit (see the analysis of the synthesis of recognition in Kant 1965). Sellars points out that “instead of coming to have a concept of something because we have noticed that sort of thing, to have the ability to notice a sort of thing is already to have the concept of that sort of thing” (1963, 176, completely italicized in original).

Perception is just one example of thinking-in-action. A broader example is skilled motor activity. I have referred above to the use of physical tools: as Polanyi and Heidegger point out, these tools become extensions of ourselves, and we can think/act through them. I have also touched on sports, where sportspersons think with their mind/bodies to solve difficult problems in real time. Only the very highest level strategies ever make it to explicit consciousness as these problems are solved, but the skilled performances of the

lower level building blocks which make these high level solutions work have required considerable explicit input for their mastery.

Yet it might be thought that while the embodiedness of perception and motor activity are relatively uncontentious and well argued, more “abstract” types of thinking are not embodied. Is all thinking embodied? My answer is “yes” and I will argue for this conclusion by considering the role of the social in the development of thinking, then the place of the body in the mind and in the physical world.

3.1252 Social Appropriation of Thinking

Picture in your mind a person thinking hard. How do you know that they are thinking? Perhaps you have pictured Rodin’s *The Thinker*, or a person striding up and down a room striking their fist on their palm, or someone staring unfocused into the distance. Baron-Cohen (1995, 105) reports on studies asking people to choose pictures that show someone thinking:

When a person's eye direction does not appear to be directed at any external object in particular - for example, when a person's eyes are directed upward and away from us but there is no external target in that part of the person's visual field - we rapidly infer that the person is thinking about something unobservable. This mentalistic inference is also drawn effortlessly by young children.

These and other physical behaviours commonly accompany thinking and, indeed, form part of the thinking; according to Baron-Cohen, we are turning our gaze, and hence our attention, inward. The behaviours are tied to the need to turn attention to the task in hand, or to provide an outlet for the committed emotional side of the thinking. The fact that we recognize such behaviours and can infer from them not only that the person in question is thinking, but also often something of the nature of their thoughts, suggests that they are more than just accompanying behaviours. They seem to form part of the act of thinking itself.

Why might such publicly expressed behaviours form part of thinking? I will argue that this is because thinking itself (or at least the large superstructural refinement of the hard-wired basis of thinking) is learned through being publicly expressed. In this, I draw on the work of Vygotsky (1981, see also §3.234):

Any function in the child’s [higher mental] development appears... first... on the social plane, then on the psychological plane. ... children... use the same forms of [thinking] behaviour in relation to themselves that others initially used in relation to them. (Quoted in Kohlberg 1987, 197)

Thinking on the social plane has to be physically expressed: both through speech and through associated body movements - especially of the eyes and (to a lesser extent) of other parts of the face. Baron-Cohen (1995, 114) lists 17 “mentalistic interpretations” which can be read from the eyes, some of which are primarily emotional, but others of which “describe ‘speech acts’ that the eyes can convey.” In the course of internalizing

this behaviour, the child not only learns (as an apprentice) how to better carry out mental activities, but also internalizes the physical behaviours that accompany, and hence become part of, the public thinking. Such psychological work reinforces the philosophical conclusions reached by George Herbert Mead, and by Wittgenstein in his famous argument against the possibility of a private language.

The insight that the natures of public argumentation and private thinking are identical is much older than Vygotsky, Mead or Wittgenstein. Billig (1987) quotes Isocrates: “the same arguments which we use in persuading others when we speak in public, we employ also when we deliberate in our thoughts”¹⁴; and Francis Bacon: “[we use the same processes] in argumentation, where we are disputing with another...[as] in meditations, when we are considering and resolving anything with ourselves”¹⁵, to similar effect. Vygotsky, Mead and Wittgenstein, however, go beyond the observation of their identity to the claim that it is precisely by participation in the public discourse that the ability to engage in private cogitation is constructed.

The child’s experience of listening to others articulate their thinking in discourse, of haltingly beginning to enter into it, of having more able others support their attempts, of gradually being able to construct more and more of the argument on their own as such “scaffolding” is gradually withdrawn, all takes place in context. Hence it is tied to the context - the physical arrangements of the participant’s bodies, the emotional and social attributes of the situation, the goals and intentions of the interlocutors, and so on - and the child will have difficulty in separating out what is essential and what is superficial. Of course, this experience is built up over many instances, and some superficial aspects of the contexts and behaviours may not be included into the child’s construction of the competencies because they are not regularly exhibited in the social situations. But many bodily cues and dispositions are so closely tied to the mental contents of certain types of thinking that it requires conscious effort to separate them off from the thinking. It is for this reason, for example, that deception is difficult - the embodied behavioural aspect of deceptive thinking can “give the game away.”

Indeed, we can go back further than this. The ability to engage in certain sorts of thinking, which are absolutely central to social interaction (and hence to ethical thought), depends critically on inferences from embodied actions to thought. I am referring here to what is commonly called the theory of mind: the attribution to others of intentional mental events. The ability to think about, and act within, the social world is constructed on this. Baron-Cohen (1995, Chapter 4) asserts that we can only build a theory of mind through certain inferences from bodily-based action to mentally-based understandings. These inferences are innate, and spring from our biological and evolutionary inheritances (see §2.4). Firstly, we infer from self-propelled motion to goals and intentions, so that,

¹⁴ Antidosis, 256

¹⁵ Of the Dignity and Enhancement of Learning, 1605

for example, we infer from a person moving towards an object that they want the object. Baron-Cohen calls this the intentionality detector. Secondly, an eye-direction detector enables us quickly and easily to detect eyes, to compute whether those eyes are directed towards or away from us, and to infer that the possessor of those eyes sees whatever their eyes are directed towards. The inferences from these two body/thought combinations can then act as input into the third module: a shared-attention mechanism, which establishes that we are having the same experience as another person, and hence allows the construction of triadic representations, when we have beliefs about the intentions of the other towards an object. Right at the basis of our ability to understand others, therefore, is our ability to read embodied behaviour as implying mental states.

In this subsection, I have suggested two ways in which publicly observable bodily behaviour inherently forms an aspect of our thinking: through the appropriation of physical mannerisms to form part of thinking; and through the hardwired inferences from bodily cues to mental events. Yet there is another way, at least as important as the construction of a theory of mind, in which our embodiment thoroughly permeates our thinking.

3.1253 The Body in the Mind

Johnson (1987) has argued in his *The Body in the Mind* that our experience of being embodied is absolutely fundamental to our ability to think and know. His argument is that our embodied existence equips us with image schemata (analogous in certain ways to emotional paradigms), which he describes as recurring dynamic patterns in perceptual interactions and motor programs. They provide fundamental ways of thinking about the world, such as up and down, near and far, force and resistance. These are knowledge of the world, but they are not propositional in any standard sense of the word; rather they are partly visual, partly kinaesthetic representations of the way we are in the world (which can also, particularly as we grow, add emotional, historical, social and linguistic dimensions in complex interactions). These provide the Background (in the terms we have been using, the situation or the context) for further, possibly more propositional, knowledge, formed by metaphorical projection: the use of patterns from one experience domain to structure another. Thus there is a physical, embodied basis that underlies and constrains all conceptual understanding, as well as our imagination.

While Johnson refers to a number of empirical studies of language use and reasoning to support his contentions, additional support comes from two quarters. The first has already been alluded to: the studies into children's understanding of everyday phenomena which show that they reason by reference to paradigm scenarios which bear a striking resemblance to image schemata. Secondly, Damasio (1996) presents evidence from neurophysiological studies that show that it makes little sense to talk of the brain, separate from the rest of the body, as the seat of the mind. The brain and the rest of the body, he

says, form an indissociable integrated organism which reacts to the environment as an ensemble. In other words, the parts of the brain that deal with data from the rest of the body do not treat it merely as input for subsequent separate processing, but involve the whole body at all stages in the activity of “reasoning” itself. The physiological operations that underlie what we call the mind are derived from this structural and functional ensemble. What Damasio labels the “somatic image” - our background sense of the state of our body - forms an essential component of all reasoning. He says: “It is not only the separation between mind and brain that is mythical: the separation between mind and body is probably just as fictional. The mind is embodied, in the full sense of the term, not just embrained” (118).

In this discussion, I have emphasized primarily only one direction of influence - from the body to the mind. Given this chapter’s focus on thinking, this is not too surprising. Yet, as Damasio’s phrase “indissociable integrated organism” indicates, we could equally well talk of the enminded body, assuming we could tolerate the barbarism of this neologism. Perhaps a better phrase is the “encultured body.” The operations and understandings of the mind radically affect the body too - in the way that it is held, the gestures it makes, the feelings it feels in response to specific circumstances and so on. Thoughts and body states are caught in an intricate feedback ensemble, which make it appropriate to talk of the organism as a person, rather than a mind in a body (Strawson 1964).

3.1254 Embodiment in a Physical World

In the previous subsection, I argued that it is not enough to recognize that the mind exists in the physical setting of the brain, but that it is permeated by its setting in the whole body. This, however, is not to go far enough. The body/mind itself does not exist in a vacuum, or even just in a social world (§3.1252). It exists in the physical world and, just as Vygotsky demonstrates that the boundaries between individual minds are blurred in social interaction (see also §3.23), our bodies are not sharply individuated from the physical world. To give but two examples, the atmosphere interpenetrates our lungs, and the tools that we find ready-to-hand extend our bodies into the physical world so that, to use Heidegger’s (1962) example, we locate our point of consciousness when we hammer in a nail at the face of the hammer (see §3.1242). Without the Vygotskian social interaction we could not think, but equally, without being immersed in, and interacting with, things themselves, we would be incapable of thought (Westphal 1998).

Westphal (1998, 26) argues that “we humans are only capable of self-conscious thought if we in fact experience a world structured by events and objects with identifiable similarities and differences”; a world with a certain sort of continuity and variety; in other words, an intelligible world. Many things make up this world: some are inanimate objects; some are animate entities. Some do and some do not use language. Information from these things that arrives at our sense organs does not carry a label to say whether it

comes from a thinking, a merely animate or an inanimate object, and we begin to categorize that information on the basis of similarities and differences. Recognition of which differences and similarities are relevant is genetically coded, as we saw in §2.4. Cosmides and Tooby claim (1994, 66):

Intuitively, we are all naive realists, experiencing the world as already parsed into objects, relationships, goals, foods, dangers, humans, words, sentences, social groups, motives, artefacts, animals, smiles, glares, relevances and salencies, the known and the obvious.... But to produce this simplified world that we effortlessly experience, a vast sea of computational problems are being silently solved, out of awareness, by a host of functionally integrated circuits. These reasoning instincts are powerful inference engines, whose automatic, non-conscious operation creates our seamless experience of the world. The sense of clarity and self-evidence they generate is so potent it is difficult to see that the computational problems they solve even exist. As a result, we incorrectly locate the computationally manufactured simplicity that we experience as a natural property of the external world - as the pristine state of nature, not requiring any explanation or research.

Since these “inference engines” have evolved through the interaction of the bodies of our ancestors in the physical world, and now allow us to situate ourselves “seamlessly” within that world (§3.124), then our thinking must inherently be coloured by the fact that we are physically embodied within this physical world.

Clark (1998, 47), as we have seen in §3.1241, argues that both philosophical and empirical evidence (largely from robotics research) points to an account of an “embodied, embedded agent... draw[ing] on the resources of mind, body, and world.” Since, on his account, much of thinking is built through physical interaction with, and scaffolding from, the physical world (see §3.126 for more detail), then the specific characteristics of the body inevitably enter into such thinking. Indeed, Clark’s inclusion of both body and world remind us that the aspects of thinking I have been discussing in these last five sections are not independent of each other. I will now look more closely at how they interrelate.

3.126 Aspectual Symbiosis

As I have already mentioned, treating the aspects of reasonableness separately, as I have done so far, risks conveying the impression that thinking can be critical, or creative, or committed and so on. This would be unfortunate, as any particular specimen of thinking must combine the different aspects. Of course, in considering a specimen of thinking, it may make sense to concentrate on one of the aspects more than the others. So, if we are looking at, say, the thinking involved in solving a mathematical problem, it might well be that the critical aspect of the thinking is of the greatest interest, while when we are considering how to convince a reluctant friend to join us in an outing, we may turn our attention more towards the committed aspect.

However, even in the broad areas of these examples, we can readily see that the other aspects are involved and may, in particular circumstances, take on a higher profile. Take the extraordinary experience I once had of teaching mathematics to a boy who was

considered very slow. I soon found out that he had not learned how to subtract if borrowing was needed, because he did not understand place value properly. Yet, when I saw him playing the darts game “301,”¹⁶ I was astounded to see him effortlessly add the scores of three darts and subtract them from the current total. Clearly, in the classroom, his obvious antipathy to mathematics, and the strange context meant that the committed and contextual aspects to his thinking ensured he could not handle a task that, from the critical point of view, was identical in form to (and easier in detail than) the dart game task. Indeed, if (as seems likely) the mere thought of maths made him feel sick, then the embodied aspect was heavily involved as well.

To turn to a completely different example, we might look at the thinking-in-action that is done by an accomplished sportsperson. The aspect of thinking that might seem to dominate in this situation is the embodied aspect. Certainly the sportsperson thinks through the body. However, if the thinking-in-action is to be good, each of the other aspects of thinking must play a role. Good sportspersons are creative players, but this does not entail just doing the unexpected, for the unexpected move to make may also be a poor move. Critical judgement is needed to decide which of the possible unusual moves is likely to be successful (and, indeed, when the expected move is the best). This critical judgement always arises in a context, so, for example, the player who will attempt one move in the dry will seek another if it is wet. Motivational factors, calling on the committed aspect of thinking, are notoriously crucial in determining the quality of sporting thinking-in-action. Many more examples of the interrelationship of the aspects of reasonableness could doubtless be advanced, but these will suffice to make the point.

I have already referred to Vygotsky’s (1962) work on the scaffolding of thought in §3.1252, and I shall return to it in more detail in §3.23 and §8.23. In those sections, however, I shall be concentrating, as Vygotsky himself did, on the scaffolding available from linguistically based social interaction. Clark (1998, 33) reminds us that scaffolding for thinking is provided by a much wider variety of structures:

Most of these strategies [to bring coherence to our thinking] involve the use of some type of external structure or ‘scaffolding’ to mold and orchestrate behavior. Obvious contenders are the immediate physical environment... and our ability to actively restructure that environment so as to better support and extend our natural problem-solving abilities. These strategies are especially important in child development. Less obvious but crucially important factors include the constraining presence of public language, culture, and institutions, the inner economy of emotional response, and various phenomena relating to group and collective intelligence. Language and culture, in particular, emerge as advanced species of external scaffolding ‘designed’ to squeeze maximum coherence and utility from fundamentally short-sighted, special-purpose, internally fragmented minds.

In Clark’s list, we find thinking is supported by both the physical and the cultural context, and by our own emotional responses, while in the last subsection, I quoted Clark on the

¹⁶ In 301, each player starts on that score and, throwing three darts in a turn, races to zero. To start requires scoring a “double,” then subsequent scores are subtracted from the remaining total. The final dart must also be a double that reduces the total to zero (i.e. if 32 is the current total, the player must throw a double 16).

place of embodiment in thinking. Since all these supports are often co-present, it is not surprising that any resultant coherent critical and/or creative thinking will also entail an admixture of these three aspects of reasonableness.

Of course, one of the most powerful features of thinking in humans is the ability to turn thinking back on itself - often referred to as metacognition (§3.1245). In taking our thoughts - that is, our mental representations - as the objects of further thinking, we can redescribe them, making them more explicit and hence more powerful (Karmiloff-Smith 1992, 15-26; Clark 1998, 207-211). Since each of the aspects of reasonableness has its own particular contribution to thinking, these contributions can be enhanced when one aspect is turned on another, as, for example, if we think creatively about our committed thinking, or critically about our contextualized thinking. Karmiloff-Smith (*ibid.*, 17) argues that

Development and learning, then, seem to take two complementary directions. On the one hand, they involve the gradual process of proceduralization (that is, rendering behavior more automatic and less accessible). On the other hand, they involve a process of 'explicitation' and increasing accessibility (that is, representing explicitly information that is implicit in the procedural representations sustaining the structure of behavior).

The latter process is carried out through meta-cognitive processes and, as she says, increases the accessibility of thoughts. These, of course, include processes as well as contents, and so metacognition enhances the transfer of competencies across contexts (see § 9.1 and §10.221 for further discussion of this point).

Donaldson (1978, Chapter 7, also 123-127) argues that disembedded thought - thought that has been "prised out of the old primitive matrix within which originally all our thinking is contained" (*ibid.*, 76) - is essential for operating fully within our society. In disembedded thought, the constraints that can be laid on thought by particularly the contextual, but also to some extent the committed and embodied, aspects of reasonableness are loosened, so that transfer of thinking can be enhanced.

I have used the metaphor of symbiosis in the heading to emphasize that the aspects are not merely contingently combined in thinking, but rather that each has its own essential element to add to the rich capability that we refer to as reasonableness. It also implies interdependence, so that while one aspect helps another, that aspect itself can turn back to assist the first, to the greater advantage of reasonableness as a whole. Thinking that lacks any one of discrimination, imagination, motivation, situational location or physical instantiation is not possible for creatures such as us.

3.13 Reasonableness and Moral Education

In this first section of Chapter Three, I have developed an account of reasonableness which is considerably richer than most conceptions of critical thinking. By shifting the emphasis away from the critical aspect (important as it is), my account redirects attention

to several important features of good thinking, particularly if it is to be considered in relation to moral education.

A multi-aspectual account is less likely to lead us towards the dubious assumption that, if only we could think clearly enough, we would all have to come to a common conclusion. If reasonableness has multiple inputs, and is always dependent on contingent factors such as the context in which it is used (which may well appear differently to different conversational partners), the commitments of those partners bring to the conversation and the way that they are embodied in the world, then reaching consensus is obviously going to be a more difficult task than if we can assume that all partners have a single-dimensional critical capacity. We shall see in §3.2 that Kant's conception of reason explicitly excludes the committed, contextual and embodied aspects of reasonableness, and that this leads him into advancing a characterization of autonomy that I will argue is seriously flawed. Attention to the empirical and theoretical literature of developmental cognitive psychology, particularly from within the Vygotskian tradition (§3.23) will provide well-documented mechanisms for the growth of reasonableness in all its aspects.

Of course, this is not to say that consensus is not ever possible, nor that it might not, in at least some circumstances, be a good idea to try to bracket the differences between the conversational partners (§7.24). We might even, as Habermas does, argue that in the ideal speech situation, such differences would be laid aside and that, hence, we must always enter discourse with the counterfactual goal of reaching consensus. In §4.1, my conception of reasonableness will be used to augment a modified Habermasian account of the pedagogic action of a teacher in a classroom, and this in turn will inform my reconceptualization of autonomy.

In Chapter Five, I will turn my attention to moral theory, and §5.2 will examine Aristotle's virtue ethics. Practical reason plays an important role in virtue ethics, and I will demonstrate that Aristotle's account of it is in considerable agreement with reasonableness, particularly in the importance it places on context and emotion. However, Aristotle does not provide us with an adequate account of the development of practical reason throughout childhood, and I will claim that this is because he has little to say about dialogical interaction. For this, I shall return to Habermas' views in more detail in §6.23, and show that, although his insistence that reason and normative correctness must always be approached dialogically does introduce a place in moral decision for commitment, context and embodiment, in the end he does not take their implications seriously enough. I will distinguish three types of rationality in Habermas' work - presuppositional, situated and ideal - and argue that we can, with some minor modifications, equate Habermas' situated rationality with reasonableness.

Subsequently, in Part III, I shall have frequent recourse to the account of reasonableness I have developed in assessing the implications for the classroom. I shall argue that the

development and use of reasonableness is a central feature of the community of ethical inquiry, and I shall explore the role it plays in the conduct of inquiry, the building of virtue, the linking of judgement to action, the avoidance of indoctrination and shallow ethical relativism, and the transfer of thinking across contexts, amongst other themes.

3.2 Autonomy

At the beginning of this chapter, I claimed that there is an interdependence between reason and autonomy. Roughly speaking, to be morally autonomous is to be an ethical agent, bearing responsibility for one's actions, and this requires reasonableness: the wherewithal to act in a way that can be supported by reasons. Alternatively, we can look at the interdependence from the other side, in recognizing that to be reasonable requires the ability to choose between reasons, an ability which is especially important in the light of a multi-aspectual account of reasonableness. But this ability to choose between reasons is itself essential to the ability to choose reasonably which action to take: which is just another way of characterizing autonomy.

Both autonomy and reasonableness are clearly implicated in education. At least in Western societies, amongst the goals of education are the formation of persons who are both autonomous and reasonable. These personal qualities are required by a democratic system that places responsibility on each individual. What is less clear is which roles reasonableness and autonomy play during schooling. Since very young children are neither fully reasonable nor fully autonomous, educators need to be clear about both what the targets of reasonableness and autonomy actually are, and how they can assist children to reach these targets.

Like much of Kant's thought, just what Kant intended by the term "autonomy" is open to much interpretation and some confusion. I will not become embroiled in close textual analysis in trying to come to terms with Kant's intentions: rather, I will survey some of the competing definitions supplied by Kant and interpretations by commentators and try to identify some of the key features, relevant to my own interests, that they share. In particular, I shall concentrate on the accounts of the nature of reason¹⁷ that underlie Kant's autonomy. Given that the two concepts are so interdependent, the considerable difference between Kant's and my account of reason will be reflected in differences in the way autonomy is seen.

In §3.21, I will firstly look at how Kant conceives of reason and the links between reason and autonomy, before pointing out the important divergences between Kant's reason and my reasonableness. One of the key problems I shall identify in Kant's work is a lack of attention to how reason develops in the growing child. I will then look in §3.22 at the

¹⁷ The words "reason" and "rationality" seem to be used pretty much interchangeably in Kant and the commentaries on Kant. I shall generally take them to be the same, and use "reason" and its derivatives in most of what follows.

interpretation of Kant as a constructivist advanced by O'Neill (1989), which offers some resources for taking the development of reason and autonomy more seriously. We shall see that there are a number of ways of talking of reason as constructed, not all of which can be applied to the developing child. Thus, I will turn in §3.23 to the evidence available from developmental psychology. Finally, §3.24 will begin on the task of reconceptualizing autonomy.

3.21 Kant on Reason and Autonomy

In Kant's eyes, humans are rational beings, "distinguished from animals by the fact that practical reason rather than instinct is the determinant of our actions" (Korsgaard 1996, 110-111). Every action, if it is not to be random, must have a cause which can be subsumed under a law, and the will is the cause of action in rational beings. But a rational will must be a free will, and a free will must be autonomous; that is, "not determined by any external force, even your own desires" (Korsgaard 1996, 25). As Kant says, autonomy is "the property the will has of being a law to itself (independently of every property belonging to objects of volition)" (Groundwork, 440, 108, quoted in Allison 1996, 134). Allison (ibid, 135) glosses the parenthetical phrase as referring to "reasons to act that... stem (even indirectly) from its needs as a sensuous being." Autonomy, says Caygill (1995, 88, quoting Kant in Groundwork, 442,46) is based on "the rational concept of perfection as a possible effect of our will" - that is, based on an ideal of perfection only, not connected in any way to the agent's particular embodied existence or experience.

In summary: for Kant, autonomy requires that emotions, desires and inclinations form no part of moral reasoning, and that the reasons involved (if action is to be autonomous) are unconnected to the embodiment or situatedness of a sensuous being in the world.

3.211 Reasonableness Compared to Kantian Reason

As we have seen above, reason and autonomy are tightly linked in Kant's account. He picks out humans as being reasonable and autonomous (in contrast to animals, which are neither). His claim is not that persons are always rational and autonomous, since we can be misled into heteronomous actions through our desires, but that humans are capable of reasonable, autonomous action when we freely choose to follow the moral law.

It can be readily seen that Kant's account of reason is markedly different from the one I developed in §3.1, and that this difference consists in the positive exclusion from reason by Kant of my final three aspects of reasonableness (the committed, contextual and embodied aspects). Since Kant's account of autonomy rests heavily on this characterization of reason, that account, too, abstracts from emotion, situatedness and embodiment. This makes moral action into a solitary activity, in which the moral agent merely considers rationally the moral law, and chooses to follow it.

However, the moral law refers to “universalizability.” It has the hypothetical form “if there are others, then I ought to treat all alike.” Moral reasoning, at least at the level of maxims and the moral law, is something that can be carried out in complete isolation from others, and even if no others actually exist. For many feminists, communitarians and neo-Aristotelians, the idea that the core of moral reasoning is the solitary, rationalistic contemplation of abstract principles seems deeply counterintuitive. They feel, and I agree with them, that if morality is about anything, it is about actual interpersonal relations and hence must arise through, and be rooted in, such relations: relations with real, flesh-and-blood, particular others. Benhabib (1992, 50) puts it like this:

The moral self is not a moral geometrician but an embodied, finite, suffering and emotive being. We are not born rational but we acquire rationality through contingent processes of socialization and identity formation. Neo-Aristotelians as well as feminist theorists have argued that we are children before we are adults, that as human children we can only survive and develop within networks of dependence with others, and that these networks of dependence constitute the “moral bonds” that continue to bind us even as moral adults.... This ‘rationalist’ bias of universalistic theories in the Kantian tradition has at least two consequences. By ignoring, or rather by abstracting away from the embedded, contingent and finite aspects of human beings, these theories are blind to the variety and richness as well as the significance of emotional and character development... This neglect of the contingent beginnings of moral personality and character also leads to a distorted vision of certain human relationships and of their *moral texture*. Universalist and proceduralist ethical theorists often confuse the moral ideal of autonomy with the vision of the self ‘as a mushroom.’¹⁸

That morality cannot be cast as the calculations of a separate, unencumbered, narrowly rational thinker has two implications: firstly, Kant’s account ignores the very important question of how it is that babies develop towards becoming reasonable and autonomous beings; and secondly, autonomy cannot be construed in a solitary fashion. Benhabib is making a claim about the importance of social bonds in the formation of persons capable of reasoning and morality. Development does not take place in a social vacuum, and the contextual, embodied and affective aspects of social settings are by no means neutral in the development of reasonableness. As Benhabib says, we are dependent on others, not merely for our physical survival, but also for our rational and moral development. I shall turn in §3.23 to a consideration of the evidence from developmental psychology for this claim.

This anticipates the thrust of Chapters Four and Six, which will look at Habermas’ neo-Kantian attempt to place discursive engagement with others at the centre of autonomy and meta-ethics. In the next section, however, I will concentrate on Benhabib’s other major point, and look a little more closely at the place of the development of reason in a Kantian account of autonomy. We will find in §3.221 that O’Neill’s attempt to take seriously the construction of reason under a Kantian account will give us pointers towards a sounder account of autonomy.

¹⁸ Benhabib’s reference to the self as a mushroom is an allusion to a passage in Hobbes (1966, 109): “Let us consider men... as if but even now sprung out of the earth, and suddenly, like mushrooms, come to full maturity, without all kind of engagement to each other.”

3.22 The Development of Reason

A traditional interpretation of Kant (see, for example, Broad 1978; Schneewind 1992; Allison 1996) is that reason seems to be something that “finite rational beings” just have: an *a priori* attribute of humans. The nature of the reason that finite rational beings have can be seen in Schneewind’s (1992, 316-318) account of Kant on the holy will.

According to Kant, a holy will is a fully rational being. Hence, the holy will recognizes the moral law (since it is the product of practical reason) and, lacking any contrary, non-rational influences, necessarily obeys the moral law. Finite rational beings also legislate the moral law for themselves. However, they sometimes do not take the right action, because they are finite: they have desires, preferences and emotions (among other factors). Finite rational beings can only be moral if they recognize that they *ought* to do what the holy will *must* do.

Both the autonomy and the reason of finite rational beings play important roles in Kant’s morality. Finite rational beings could acquire knowledge of the moral law in two ways: they could be told what the moral law is and that it holds for them; or they could have legislated the moral law for themselves through their reason. In the first case, the decision to follow the moral law would be made heteronomously and thus could not be a moral decision. Hence for Kant, only autonomous, reasoning beings can understand morality at all (to be able to know what the holy will must do, and therefore what one ought to do). Humans, as finite rational beings, therefore have perfect rationality tempered by their finitude (desires etc), rather than being finite in their rationality.¹⁹

¹⁹ Cherniak’s (1986) work on minimal rationality calls into question the supposition that the rationality of a finite being would be just a more limited mirror of the rationality of an unlimited being. For example, he argues that it is rational for a finite being to accept that some of its beliefs will be false, given the time and effort required to root out such false beliefs when compared to the penalty for holding them. Such an acceptance would clearly be irrational in an unlimited being, for whom time and effort are no problem. Equally, it would be irrational for a limited being to make even a small fraction of the sound and feasible inferences from its belief set, given the paralysis that would result, but not so for the holy will. The desires and preferences of the limited rational being enter inherently into this sort of rationality. In deciding in which areas of our knowledge it is worth investing a lot of effort to track down false beliefs, or when it is essential to draw more inferences, the agent must make judgements of importance and immediacy. These judgements can only be based upon what matters most to the agent - something that is governed by desires and preferences. A holy will has no desires and preferences, and no need to make such decisions. So even the core rules of rationality differ between limited rational beings and holy wills. Cherniak calls into question even the possibility of defining the rationality of a holy will: “For we have found that there seem to be remarkably few *a priori* constraint on a rational agent’s deductive abilities. An agent must only be an adequate logician; beyond generalizations about human psychology, there seems to be no transcendently ‘right’ kind of logician” (48). Formal reasoning systems are not an idealization of the actual techniques used by humans who are acknowledged to be thinking and acting reasonably. The finite individual’s reason is not that of the transcendental being. It seems that Kant has misconceived the relationship between a “finite rational being” and a holy will. The two cannot be, as he requires, beings with equivalent reason, but with the former limited by alien desires and preferences. Kant would have the adjective “finite” modify “being,” not “rational,” yet humans must have limitations that apply to their rationality itself. Kant’s account assumes an objective reason, that is identically available to all. Humans, with their desires and preferences, cannot have such reason.

What is not clear in this account is the place of children. Are all children (as humans) both autonomous and possessors of innate reason? Kant makes it clear that even if finite rational beings are autonomous, that does not mean that they always act autonomously or rationally - they are often led by their limitations to act heteronomously. The limitations most commonly cited are desires and preferences, but other aspects, such as lack of experience or possibly continuing brain development, may mean that, although children also have innate reason and autonomy, they are limited in their ability to access or apply them. Thus the inability of young children to reason correctly and act autonomously can be explained, and does not count as a counterexample to Kant's thesis.

This account would maintain reason and autonomy as attributes of all humans, even children. Development then would be a case of overcoming the limitations that block reason and autonomy. Yet such an account seems to be at odds with our experience of child development and education, where it seems that children are actually learning how to reason, rather than how to avoid limitations. I shall turn in §3.23 to empirical evidence that supports this claim. Before I do, however, I shall turn my attention to a more recent interpretation of Kant that claims he supported a constructivist view of reason.

3.221 A Constructivist Version of Kant

O'Neill (1989) has defended an unorthodox anti-foundationalist and constructivist account of Kant's practical philosophy. Her interpretation is at odds with most commentators, who have taken Kant to base his practical philosophy on *a priori* reason, as we have seen above. However, O'Neill points out that Kant's overall project is usually taken as (successfully) demonstrating the inadequacy of foundationalist metaphysics, while having surreptitious recourse to it in his practical philosophy. She takes seriously the possibility that Kant held to anti-foundationalist views throughout, and develops her account on that basis. Her interpretation seems to rest heavily on the following quotation, which appears three times in the first three chapters of her book (pp. 15, 37, 57).²⁰

Reason must in all its undertakings subject itself to criticism... Nothing is so important through its usefulness, nothing so sacred, that it may be exempted from this searching examination, which knows no respect for persons. Reason depends on this freedom for its very existence. For reason has no dictatorial authority; its verdict is always simply the agreement of free citizens, of whom each one must be permitted to express, without let or hindrance, his objection or even his veto. (*Critique of Pure Reason*, A738/B766)

Reason is no longer *a priori*, a given that we can build on. Rather, reason is a construction in which all (or, in Kant's day at least, all educated men) can and indeed should help. In further support of her claim, O'Neill refers to Kant's writings on the history of reason, and to the very idea of a critique of reason, for the word "critique"

²⁰ This passage anticipates the approach of Habermas, who has drawn heavily on the Kantian tradition: see §6.23. As in Habermas' account, the existence of other persons is central to O'Neill's reconstruction of Kant.

implies a finding of weaknesses and the attempt to improve (“Critique of reason is only possible if we think of critique as recursive and reason as constructed rather than imposed,” p 27). In a claim that echoes my position in Chapter Two, she says that Kant holds that the roots of reason lie in the “cunning of nature,” which provides “just enough [reason] for the most pressing needs,” before social and cultural development takes over (quotes are Kant’s, from O’Neill 1989, 39). Kant was optimistic that the task of finding the most pure, best reason would not take long - “before the end of the present century” he wrote in 1780 (quoted in di Giovanni 1992).

I think there are pressing problems with this interpretation of Kant. There are a number of ways in which we can interpret the phrase “the construction of reason.” Firstly, we need to make a distinction between the construction of reason itself and the construction of an adequate account of reason, and I think that the passages to which O’Neill appeals to support her constructivist account can be best read to support the latter. As I have pointed out in §3.124, we are able to do many things without being able to give anything like a full account of how we do them. If reason is innate, then while that means all humans are “rational beings,” it does not imply that all humans can describe and explain their reason. Once we realise this, we can see that Kant, in the passages quoted by O’Neill to support her interpretation, is merely arguing that it is our human *account* of reason that needs to be constructed. The task for which he calls upon others to assist him is this descriptive and explanatory task, not the construction of reason itself, which he takes to be in some sense innate.

Even if I am wrong about this, there are further ambiguities. The construction of reason itself (rather than an account of it) can be considered historically, as a production of humanity, or personally, as an individual attainment.²¹ Again, the Kantian passages upon which O’Neill relies support the former interpretation, seeming to say little about the growth of reason *in persons*. Yet, if the capacity to reason and act autonomously was constructed historically in humans (as I think we must agree it was), then there are some problems with O’Neill’s account of the interactions between reason and autonomy.

Reason, says O’Neill in one place, is based on freedom and autonomy.

This gap [between reason and autonomy] is to be closed not by establishing that human beings are rational, and then proving that they are free and hence also (given the analytic argument) autonomous. Kant’s strategy is the reverse.... Only autonomous, self-disciplining beings can act on principles that we have grounds to call principles of reasoning. Reason has no transcendent authority; it can only be vindicated by critique, and critique itself is at bottom no more than the practice of autonomy in thinking. Autonomy does not presuppose but rather constitutes the principles of reason and their authority. (57)

²¹ I do not wish to imply by the phrases “personal construction” or “individual construction” that the construction of reason is something that people can do on their own, in social isolation. Indeed, I will argue just the opposite. These phrases are meant to draw a distinction between reason in Popper’s Third World sense (the joint historical product of many thinkers - here called historical construction) and reason as it becomes constructed in an individual’s thinking repertoire.

Now, there are problems seeing how the concept of self-legislation can lead to the myriad of skills, abilities, capacities and capabilities that we call reason, even if reason is restricted in the Kantian way to rationality. However, I shall pass over those, for only a few pages earlier, O'Neill seems to be saying that autonomy is based on reason:

We can begin to see why, if autonomous action is to be independent of everything 'alien,' it must be action determined by reason. Independence from 'alien' causes, taken strictly, must be independence from contingent or variable events, including those that are intimately part of agents. Such events depend on the play of natural forces or (possibly) on random occurrences, and so ultimately on indisputably 'alien' causes. Maxims of autonomous action, by contrast, must hold equally for all rational agents, whatever their peculiar contingent and variable characteristics, and so must be universalizable. Their authority, if they have any, cannot derive from any contingency of human life but only from the requirements of reason, whatever those may be. (54)

In this passage, O'Neill seems to be saying that autonomous action can only be action determined by reason and, further, that maxims of autonomous action must arise from the requirements of reason. In two (connected) ways, then, reason predates autonomy. For without the underpinnings of reason, no action can be autonomous.

This apparent contradiction could be resolved by recourse to Allison's Reciprocity Thesis; that autonomy and reason reciprocally imply each other (Allison 1996, Chapter 9). This account might be sustainable on an *a priori* account of reason, where reason and autonomy are all or nothing affairs. Difficulties arise if reason is constructed; particularly if it is not just constructed jointly over historical time by adult thinkers, but also individually over a lifetime by embodied growing children. Autonomy can only support reason if it is already present, and vice versa. Of course, we might consider some sort of cyclical bootstrapping mechanism (and this is indeed similar to the account that I will subsequently develop), but this would seem to have the consequence that children could not be morally accountable until both reason and autonomy have been constructed, for it is not until then that they can self-legislate the moral law. Moral development reduces to a one-off moment of self-legislation. Indeed, it is a clear consequence of Kohlberg's Kantian-based account of moral development that children *are* pre-moral until this "moment of enlightenment," as I will discuss in §5.1.

O'Neill's further claim in this passage is that Kant needs reason to be a single capacity, shared by all rational beings. Only then can it be said that a decision made on the basis of reason alone is autonomous. This is because if reason differed from agent to agent, then what reason commanded in one agent would be contingently different from what it commanded in another agent, and hence their actions would be heteronomous.

Yet, if O'Neill is committed to the individual construction of reason, the current construction of reasoning available to an individual *is* dependent on contingent factors, and Kant's argument shows us that the individual is heteronomous. Constructions can be more or less completed and, further, the construction need not proceed following an identical plan in all cases, as it involves input from five different aspects (§3.12). The

only escape from this conclusion seems to be to understand reason and autonomy as relatively minimal constructions, a subset of the agent's total construction, so that it may be that all (normal) adult humans have completed the task. If all human agents of the requisite age had already constructed the required subsets of reason and autonomy, they may be able to impose the moral law on themselves. Just how an agent would know precisely which subset of reason ought to be accessed in order to do the self-legislating is not clear.

A minimalist position on reason in itself is problematic, as the argument that Kant mounts in order to support his contention that all humans, *qua* rational and autonomous beings, impose the moral law upon themselves is long and complex; highly unlikely to be able to be followed, let alone autonomously initiated, by the vast majority of those whom he claims have done so. He does not claim that in order to impose the moral law on oneself, it is necessary to work through this reasoning: it is merely necessary to be autonomous and capable of rationality. Yet, exactly how this self-legislation is achieved without following his argument is unclear.

Of course, we can ask at this point which we should abandon: the construction of reason or the Kantian account. I shall argue that, as we have good grounds for believing that reason is personally constructed, then it is Kantian notions of reason and autonomy which need to be abandoned. I have already developed an account of reasonableness in §3.1 that takes account of aspects of reasonableness that are not merely ignored, but positively excluded, by Kant. It is time now to look at the empirical evidence about how such reasonableness could be constructed by developing children. Then, equipped with a better characterization of reasonableness and its construction, I will turn again to autonomy.

3.23 Evidence from Experimental Developmental Psychology

To say, as I said in Chapter Two, that certain features of the brain and its operation are present at birth is not to say that human reason is already fully formed. On the contrary, there is still much work to do before a child can be said to be reasonable. This is the social and experiential superstructure of reason and, as I stated in the previous subsection, it further ensures that reason is not universally identical. In this section, I will survey some of the research into how children's reasonableness grows and develops. The main thrust of this section will be that the major mechanism for the development of reasonableness is Vygotskian learning.

The Vygotskian learning framework is a version of genetic epistemology, originated by Piaget. Vygotsky's major work, *Thought and Language* (1962), begins with an explicit reaction to Piaget. In its broad sense (that is, not used purely as a name for Piaget's research program), genetic epistemology is concerned with the origins and development of both children's knowledge itself and children's ability to acquire knowledge as they

mature and grow. Since almost all modern work on cognitive developmental psychology arises either as a continuation of, or in reaction to, Piaget's claims, I will first outline the key features of his account. The summary of his work that follows depends heavily on Donaldson (1978, 129-146).

I referred above to the superstructure of reasonableness. There are two aspects to this superstructure: the social, which relies on interaction with other humans, and the experiential, which relies on an individual's solitary attempts to make sense of the world. Clearly, such a division is somewhat arbitrary, as social interaction is mediated through the world, and experiential exploration is often carried out socially. Nevertheless, the division may stand for the time being in order to make a few points.

Experimental study of the experiential exploration of the world, and its connection to the construction of reason, has been most famously carried out by Piaget. Although Piaget took some interest in the development of specific concepts (or bits of knowledge), particularly in his 1929 *The Child's Conception of the World* (Solomon 1994, 2-3), the main thrust of his work was to explicate how children gained the ability to gain knowledge. In other words, he was more interested in the cognitive strategies available to children at different ages, and how they came about, than he was in studying what children learn.

Piaget characterizes a human as a self-regulating system that builds models of the world and how it can be interpreted. When such a model is adequate to experience, the person is in equilibrium. Further activity, however, will throw up experiences that do not neatly correspond to that model. This dissonance can result in action of one of two kinds: either the person can see a way to incorporate the new experience into the broad models - knowledge webs or cognitive tools - already being used (*assimilation*); or they will realise that it cannot be fitted in without changing the model (*accommodation*). In most cases, both processes are at work together.

One of Piaget's key claims is that part of the model is the inventory of intellectual tools used to deal with the experiences: that is, reason. When the presently held set of tools is seen to be inadequate to deal with experience any more, then reason can be reconstructed through accommodation. Piaget claims that when these reconstructions take place, the child enters a new *stage*, where new reasoning tools become available and are potentially useable across the whole range of the child's interaction with the world. Thus, Piaget's theory is often known as a Stage Theory. Examples of such tools include reversibility and decentration, but the details of Piaget's stages are not important here.

Piaget's theory has increasingly come under criticism in a number of respects. It is as well to be careful here, however, as the theory was constructed over more than 50 years, and there are apparent internal contradictions in his writings. Further, there is good

evidence that his theory was being adapted towards the end of his life to accommodate at least some of the objections (Fosnot 1993). The four points in Piagetian theory (as generally understood) that I think are the weakest are:

1. the monolithic stage theory implies that all reasoning is domain-general. Current evidence is that reasoning is context sensitive. This question of transfer across domains was addressed in §3.1245, and I shall return to in §10.221, so I shall merely note here that Karmiloff-Smith's (1992, 11) judgement is that "Piaget's... arguments for across-the-board stages no longer constitute a viable developmental framework."
2. the role of social interaction is downplayed in comparison to the role of solitary thought. This point is perhaps the most controversial, for it is relatively easy to find passages in Piaget's writings that support either the contention that social interaction is of minimal importance, or that it is of great importance.
3. language is confined to the role of representation and reporting, never given a role as a source of thought.
4. learning is said to be different from, and to follow after, development. Thus, education becomes largely a matter of watching for cognitive development before introducing new, more demanding material, rather than creating the conditions to encourage development through new learning.

3.231 Social Interaction

Donaldson (1978, 145) points out the interaction between numbers two and four:

For [Piaget] 'learning' is by no means synonymous with 'development'. Rather, he tends to equate 'learning' with the acquisition of knowledge from some external source - that is, he contrasts it with acquisition as the result of one's own activities.

Solomon (1994, 3) quotes Piaget writing dismissively in the foreword to a book by his collaborators (Inhelder, Sinclair et al. 1974, who had tried to see if learning could accelerate cognitive development): "bits of learning are not development." She notes that Piaget (not surprisingly, given this opinion) was not very interested in teaching, and records that he rejected attempts by Bruner to interest him in educational work.

Even in as social a field as moral development, Piaget (1977) thought that the actual development of a cognitively different moral capacity is the result of "intelligence working on moral rules as on all other data by generalizing them and differentiating between them" - a personal and internal action. He tied such development to a domain-general entry into a higher stage, where reciprocity has developed: "Autonomy therefore appears only with reciprocity, when mutual respect is strong enough to make the individual *feel from within* the desire to treat others as he himself would wish to be

treated” (both quotes from 189, my italics). It is worth noting the similarities with Kant’s account of the moral law - treating others as one wishes oneself to be treated - as something that must be imposed by a solitary thinker from within. This is not coincidence, for Piaget’s work draws explicitly on Kantian philosophy, and this Kantian lineage reappears in Kohlberg’s neo-Piagetian account of moral development (see §4.2).

Karmiloff-Smith (1992, 17-26) presents a neo-Piagetian account of development that renounces global stages. Nevertheless, her model of “representational redescription” is largely internally driven. She claims that knowledge can be represented in one of four formats. The first is an intuitive format, where information is encoded in procedural form, not available for any use outside the module. Subsequently, children feel a drive to understand their knowledge, and redescribe the representation into successively more explicit formats, available more generally. While this model accounts for domain specificity, and later domain generality, of knowledge better than Piaget’s model, Karmiloff-Smith still pays little attention to the social element in the acquisition of knowledge and cognitive competency. It rates a single mention in her book - in relation to social development. To be fair, she claims that this is “partly because a number of developmental theories... have in my view given too much weight to social interaction at the cost of neglecting important endogenous factors” (ibid., 122).

Vygotsky (1962, 23) puts his finger on the second problem when he says that “the developmental uniformities established by Piaget apply to the given milieu, under the conditions of Piaget’s study. They are not laws of nature but are historically and socially determined.” He continues by pinpointing the key respect in which the conditions of Piaget’s studies were limited: they did not involve studying children in social interaction. In such a milieu, Vygotsky (1981) finds a different process of cognitive development:

Any function in the child’s [higher mental] development appears... first... on the social plane, then on the psychological plane. ... children... use the same forms of [thinking] behaviour in relation to themselves that others initially used in relation to them. (quoted in Kohlberg and Wertsch 1987, 197)

In *Thought and Language*, he puts this in perhaps simpler terms: “What the child can do in cooperation today, he can do alone tomorrow” (Vygotsky 1962, 104). Like Piaget, he stresses the importance of challenging the thought of children as a means to improve it. But while for Piaget the challenges come, especially before about age 12, from the child’s interaction with the world, Vygotsky sees a much greater role for language and social interaction. Forman and Cazden (1985) expand on Vygotsky’s thesis as follows:

Each child learns to use speech to guide the actions of his or her partner and, in turn, to be guided by the partner’s speech. Exposure to this form of social regulation can enable children to master difficult problems together before they are capable of solving them alone. More importantly, experience with social forms of regulation can provide children with just the tools they need to master problems on their own. ... When they can apply this social understanding to themselves, they can then solve, independently, those tasks that they had previously been able to solve only with assistance.

Piaget, whose clinical interviews were conducted almost exclusively with children on their own, misses (or possibly underemphasizes) two essential points: that children, in verbal interaction with others, are capable of participating in better thinking than a single one of them alone can manage; and that children construct their reason from their experience of the thought processes of others as expressed in discussion and social interaction (and not simply through development and physical experiences). Empirical evidence to support these conclusions continues to mount (Donaldson 1978). Tizard and Hughes (1984, 126), on the basis of transcripts of children's conversations, point out that "by the age of three or four... dialogue is as important as physical exploration" in the construction of reason.

After presenting his own empirical evidence to support the view that young children learn how to reason from a contradiction in this way, Lawson (1990, 550) addresses the educational implications: if "one views development as a consequence of the internalization of the patterns of social (verbal) discourse, then a failure of the classroom to engage students in such discourse would deprive them of opportunity to develop the reasoning patterns in question."

3.232 The Role of Language

Rather than giving language merely a representational and reportative role as Piaget does, Vygotsky argues for a close connection between thought and language that, once established in early childhood, leads to a greater empowerment of both. The communicative role of language, he argues, is prior to its representational role. The phenomenon of self-directed speech - the habit young children have of talking to themselves as they accomplish tasks - is interpreted differently by Piaget and Vygotsky. The former takes it as a sign of the egocentrism of the young child and its disappearance as a waning of egocentrism: hence, he calls it egocentric speech. Vygotsky interprets it as a reproduction by the child of the sort of commentary that others, usually adults, have used to accompany shared tasks, and the disappearance as the transformation into "inner speech" or verbalized thought. Language and thought are thus intimately intertwined. Later, further internalization of dialogical interchange is the basis for the development of reasoning, as the child incorporates reasoning capabilities used in the social setting.

Mead, working independently in the USA, came to similar conclusions to Vygotsky. Kohlberg and Wertsch (1987) summarize Mead's findings: "The basic mechanism that makes the socialization of thought possible... is the process of dialogue" (209). "Cognitive self-guidance speech and hence inner speech or verbal thought... must develop out of a prior dialogue mode" (214). Sellars goes even further, claiming that "the ability to have thoughts is acquired in the process of acquiring overt speech and that only after overt speech is well established, can 'inner speech' occur without its overt culmination" (1963, 188). In this, he seems to deny that any thought takes place at all

before the onset of language acquisition, but it is as well to recall that Sellars restricts “thought” to propositional thought, in a way that neither Vygotsky nor Mead do.

3.233 Learning and Development

According to Vygotsky, if we study the development of cognitive capacities in social settings, we find a very different relationship between learning and development. He says that, “since instruction given in one area can transform and reorganise other areas of child thought, it may not only follow maturing... but also precede it and further its progress” (Vygotsky 1962, 96). In other words, learning does not just wait on cognitive development, as Piaget asserts, it also often leads to development. Learning and using language plays a central part in this process:

Thought development is determined by language, i.e., by the linguistic tools of thought and by the sociocultural experience of the child.... The child’s intellectual growth is contingent on his mastering the social means of thought, that is, language.... [When the child learns to talk] *the nature of the development itself changes*, from biological to sociohistorical. Verbal thought is not an innate, natural form of behaviour but is determined by a historical-cultural process and has specific properties and laws that cannot be found in the natural forms of thought and speech. (*ibid.*, 51, original italics)

Kohlberg and Wertsch (1987, 197), in a thoughtful comparison of Piaget and Vygotsky, recognize that the latter “emphasised that the developmental analysis of higher mental processes must begin with an examination of their linguistically mediated social origins.” If we want to know how the ability to think develops and improves throughout childhood, we need to concentrate on the social linguistic interchanges between children and others.

The key Vygotskian insight is that children can engage verbally in a social group in forms of reasoning that are not yet available to them individually, and that it is precisely this engagement that enables children eventually to use those forms of thought. Vygotsky uses the term “zone of proximal development” to label this conceptual and reasoning space in which children can operate with help from a group, but are not capable of operating in on their own.

Thus, there is solid empirical and theoretical support from developmental psychology for the view that reason is constructed. Piaget provides evidence for the role of assimilation and, especially, accommodation in interaction with the world in this process, while Vygotsky emphasizes that children can, in a social situation, work beyond their individual competencies, and that the subsequent internalization of these competencies means that the construction of reason involves a primary role for social dialogue and language.

3.234 Vygotsky and the Aspects of Reasonableness

So far I have referred to cognitive capacities in general, claiming that Vygotskian learning applies to the development of reasonableness. Given the multi-aspectual nature of my account of reasonableness, it is important to make clear that Vygotsky does not construe rationality narrowly, encompassing merely what I have called the critical aspect of

thinking. My focus below will be on the aspects of reasonableness, though I shall also appeal to a Vygotskian analysis of the acquisition of virtue in §5.22. Vygotsky says:

The fruitfulness of our method may be demonstrated also in other questions concerning the relations between consciousness as a whole and its parts.... We have in mind the relation between intellect and affect. Their separation as subjects of study is a major weakness of traditional psychology since it makes the thought process appear as an autonomous flow of "thoughts thinking themselves," segregated from the fullness of life, from the personal needs and interests, the inclinations and impulses of the thinker. Such segregated thought must be viewed as either a meaningless epiphenomenon incapable of changing anything in the life or conduct of a person, or else as some kind of primeval force exerting an influence on personal life in an inexplicable, mysterious way... By the same token, the old approach precludes any fruitful study of the reverse process, the influence of thought on affect and volition... Every idea contains a transmuted affective attitude towards the bit of reality to which it refers. (1962, 8)

In this, we can see that Vygotsky gives an account of what I have called the committed aspect of thinking, together with the contextual aspect. Indeed, as he argues here, it is the fullness of life - the context in which we live - that links the emotions to the contents and processes of thought. But there is also an essential place in Vygotsky's scheme for the creative aspect of thought, again linked to the context in which thoughts form and evolve:

Concept creation is a creative, not a mechanical, passive process... a concept emerges and takes shape in the course of a complex operation aimed at the solution of some problem... the mere presence of external conditions favoring a mechanical linking of word and object does not suffice to produce a concept. (54)

In arguing that concept formation is a complex and ongoing process, Vygotsky provides supportive grounds for Johnson's (1993, 90-99) prototype account of concepts, an account that I shall return to, particularly in §6.2322 and §7.24.

a concept is more than the sum of certain associative bonds formed by memory, more than a mental habit; it is a complex and genuine act of thought that cannot be taught by drilling but can be accomplished only when the child's mental development itself has reached the requisite level. At any age, a concept embodied in a word represents an act of generalization. But word meanings evolve. When a new word has been learned by the child, its development is barely starting... as the child's intellect develops, it is replaced by generalizations of a higher and higher type - a process that in the end leads to the formation of true concepts. ... Practical experience also shows that the direct teaching of concepts is impossible and fruitless. (82)

Concepts form in multiple contexts, evolve with age and use, take on characteristics rooted in the settings in which they form and metamorphose, generalizing from those contexts but retaining the marks of their genesis. Thus concepts are key carriers of the contextual aspect of reasonableness.

3.235 Development, Reasonableness and Autonomy

In this section, I have outlined the evidence that reasonableness is constructed in social interaction with others. This, especially when taken in conjunction with the multi-aspectual account of reasonableness developed in §3.1 and reinforced above, demonstrates that Kant's account of reason as rationality, shared by all humans, cannot be the basis for an account of autonomy. I will now turn to a first pass at taking account of

social interaction in characterizing autonomy. Following this section, in Chapter Four, I will look at Habermas' neo-Kantian theory of communicative action that emphasizes the importance of social interaction through discourse, considering its implications for the classroom. This, we shall find, can overcome the lack of a developmental element in Kant's account of autonomy, but will itself need to be modified somewhat in §4.2, where I shall present my full version of autonomy.

3.24 A Reconceptualization of Autonomy

3.241 Thinking for Oneself

It is a commonplace that thinking for oneself is a good thing. Thinking for oneself is taken to be a hallmark of autonomy. Yet it is far from clear that those who endorse "thinking for oneself" as an aim have thought through properly what it means. Let's look at an example.

Jackson (1994), in a chapter entitled *Approaches to Learning Ethics* tells a story taken from Arnold (1888, 205) about Mary Shelley. Upon being given the advice to "send [her son] somewhere where they will teach him to think for himself," Shelley replied "Oh, my God! teach him rather to think like other people!" Jackson comments:

Thinking for oneself rather than thinking like other people, that is the question. Most ethics educators will say they want students to think for themselves. Independence and autonomy are often cited as moral imperatives... If ethics educators value autonomy, then let us take advantage of the research into teaching and learning... to advance towards it.
(179)

I am not saying that thinking for oneself is worthless and that people ought rather to thoughtlessly parrot others. If this was Mary Shelley's meaning, then I don't agree with her.²² However, there are three points to be made. The first is that the dichotomy posed by Jackson was not thinking for oneself vs thoughtlessly espousing the thoughts of others, but rather thinking for oneself vs thinking like others. The second is that there is no mention of the quality of thinking, just who is doing it. Thirdly, it is impossible to think for oneself about everything that one is to accept as knowledge.

3.2411 Thinking for Oneself vs Thinking Like Others

One is not autonomous if one accepts, without thought, the thinking others have done. But there is no necessary dichotomy between doing one's own thinking (considering, weighing, exploring etc) about an issue, and coming to the same conclusion that everyone else who has thought carefully about that issue has come to. Indeed, if many have thought about an issue and come to a particular conclusion, then the fact that an individual, based on their own thinking, comes up with another conclusion is a *prima facie* case for rejecting that conclusion. Of course, if cogent reasons can be given for

²² Since Mary Shelley, author of *Frankenstein*, was the daughter of Mary Wollstonecraft (*A Vindication of the Rights of Women*) and William Godwin (*Political Justice*) and wife of poet Percy Bysshe Shelley, all (for the time) notorious free-thinkers, she may have had good reason for her remark.

accepting it and rejecting the consensus, this individual thinking will be exceptionally valuable, but it is not the mere fact that it has been thought about individually that makes it so. Original thinking that “breaks the mold” is always valuable, but not all idiosyncratic thinking is valuable.

3.2412 The Quality of Thinking

This is where the second point bites. Thinking for oneself is of little value unless the thinking is of sufficient quality. Quality thinking is not something that takes place according to idiosyncratic rules or processes. Rather, it is thinking that can be judged to be good against certain criteria. These criteria must be public, and the judging potentially so. Thus, in order to think *well* for oneself, one must, by and large, think like others who also think well. This is not to say that all good thinkers think the same thoughts: we can make a process/content distinction here. Good thinking utilizes good, standard processes (which may be algorithms or heuristics), but can be applied to new content, or old content in a novel way, to come up with new conclusions. Equally, an individual can string together thought processes which are common in an idiosyncratic yet valid way. The generative rules that cover logic and language allow creativity of this sort within rule bound behaviour - rules which are publicly available to all.

We have to be careful here. If we take the lessons of Cherniak (1986) to heart, then there is no ideal rational procedure mapped out in advance for dealing with a particular information set (drawing all possible inferences from that set, for example). To attempt to proceed according to some such notion of pure rationality would be a waste of resources and hence irrational. Good thinking is an art, not merely a matter of following strict algorithms. Like art, it does require mastery of techniques which may themselves be largely algorithmic, but it also requires much more - for example, nuanced and situated judgement about which algorithms, if any, to apply in which contexts.

Nothing said above is meant to rule out the possibility of individuals discovering for themselves, for example, new ways of constructing arguments in logic, or a valuable new heuristic to use in a certain context. Such breakthroughs, if entirely new, are very rare - and even so would need to be assessed against public criteria. This is to say that the rules and criteria for good thinking are themselves not fixed. They can become the content of thinking, and they may be able to be improved, if the argument for a better rule or criterion can be sustained by good thinking. The content/process distinction is useful, but it is not set in concrete. Further, this process of validation is public and consensual, not a matter of thinking for oneself.

Further, it has to be recognized that thinking is not merely about facts. We apply our thinking to our values, our personal, social and cultural beliefs as well as our factual knowledge, and we do it from within the context in which we are situated. Even factual knowledge is not universally shared, not even amongst those who are in close proximity.

Given this variety in starting points, it is not surprising that two equally competent thinkers, perhaps even using the same thought processes, will not necessarily come to the same conclusions.

3.2413 Testimony

The third problem with lauding thinking for oneself was hinted at when I characterized thinking for oneself as “considering, weighing, exploring etc.” If we are not thoughtlessly to take on the thoughts of others, does this imply that we need to carefully consider, weigh and explore every thought that others express to us? Such a model of thinking for oneself would be impossible. Given the large proportion of our thoughts which we take on trust from others with whom we interact, such a task would be paralysing. Rather than needing to rethink and validate all knowledge offered by others, thinking for oneself is often more about evaluating the conditions of trust. Is this person generally trustworthy? Are they likely to be trustworthy on this particular point? Under what circumstances should I reopen the question of validity for knowledge that I took on trust earlier? We need to be wary of asserting that thinking for oneself requires global scepticism of others, cutting ourselves off from reliance upon them (Code 1991). In many cases, we must think like others in the absence of any good reason to do otherwise.

This point cuts deepest when we consider how it is that we learn what good thinking is, rather than just considering content knowledge. Young children are as yet incapable of evaluating the adequacy of the procedures of reasoning that are used by more capable others in their social setting. They must take them on trust if they are to become capable of reasoning at all. Eventually, as they are becoming more competent reasoners, they will be able to reflexively turn the tools of reason back onto themselves, but this cannot be done at every stage.

3.242 Autonomy, Heteronomy and Dependence

In considering the three points above, I argued that there is an important sense in which thinking for oneself *is* thinking like other people, provided the other people concerned are good thinkers. Due to the support that the products of thinking receive through the collaboration of others, we have to be especially vigilant of our own thinking when it *doesn't* issue in generally accepted conclusions. To do this requires us to be clear that our own thinking is of the highest possible quality, but the only test of this is a public one, against public criteria. Indeed, to become capable of such reasoning requires a dependence on the testimony of others, not only for content, but also with regard to process.

Autonomy would certainly require thinking for oneself, so we need to be careful in our characterization of exactly what this means. I have argued that it does not mean that we must cut ourselves off from all influence by, and dependence on, others. To do so, I have

argued, would be to risk either never being able to become reasonable in the first place, or falling into unreasonableness. Such considerations indicate that a distinction can be drawn between heteronomy - being governed by others - and dependence on others. It is only through dependence on others that we can learn to think *well* for ourselves, and only by becoming able to think well for ourselves that we can approach autonomy.

The Kantian conception of autonomy excludes any hint of dependence, not only on others, but also on one's own desires, preferences and emotions. This view arises out of a tradition of analysis that takes as its starting point the fully grown up, competent person, usually male, whose rationality is identical to that of others, being in no way contingent on a developmental life history amongst others. Thus Hobbes (1966, 109) says:

Let us consider men... as if but even now sprung out of the earth, and suddenly, like mushrooms, come to full maturity, without all kind of engagement to each other

That such an idealization of persons is a sound philosophical launching pad continues to be asserted in the neo-Kantian work of the present, such as that given by Rawls (1973, 137-9) in his description of persons in the original position:

First of all, no one knows his place in society, his class position or status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like. Nor, again, does anyone know his conception of the good, the particulars of his rational plan for life, or even the special features of his psychology such as his aversion to risk or liability to optimism or pessimism.... As far as possible, then, the only particular facts that the parties know is that their society is subject to the circumstances of justice and whatever this implies.... It is clear that since the differences among the parties are unknown to them, and everyone is equally rational and similarly situated, each is convinced by the same arguments. Therefore we can view the choice in the original position from the standpoint of one person selected at random.

Kant, Hobbes and Rawls, as we have seen, rest their philosophical arguments on solitary individuals, abstracted very much from ordinary life. This position can be traced back to Descartes (1984), sitting alone in his armchair, pondering what he could believe in and coming to the conclusion that the only certainty lay inside his own head. This Cartesian legacy, and the characterization of autonomy it spawned, have increasingly come under siege.

3.3 Towards a New Understanding of Autonomy

Many post-modernists, communitarians and feminists have attacked autonomy, casting it as an Enlightenment concept, a notion that denies human connectedness, an idealization of an educated Western male that ignores culture and excludes most of humanity. The feminist attack, for example, claims that autonomy has been conceptualized in a way that privileges males above females. Thus, Young (1997, 124) discusses:

personal autonomy, a sense of self confidence, and inner direction, as well as the ability to be reflective, not swayed by immediate impulse or blind emotion in the making of political argument. Paradoxically, such autonomy and personal independence is thought to require the loving attention of particularist mothers who devote themselves to fostering this sense of self in their children. Attentive love disqualifies the nurturers of the individuality and autonomy of citizens from the exercise of citizenship, however, because

the character of mothers tends to be emotional and oriented to particular needs and interests instead of to the general good.

This is how Haste (1996, 51), speaking on behalf of communitarians, puts it:

People simply are not solitary beings capable of 'autonomous reasoning' from behind a veil of ignorance, or in a state of suspended objectivity from one's cultural context. It is meaningless to talk of people 'stepping outside' or 'transcending' their culture and time, however 'rational' they try to be. Therefore it is pointless to make autonomy an ideal either as a personal quality or as a form of reasoning.

In Young's and Haste's comments, it is notable that precisely the neglected aspects of reasonableness that are highlighted. Young refers to love and nurture, both clearly connected to the committed aspect of reasonableness, while Haste refers to cultural context and time, factors that are covered by the contextual aspect of reasonableness.

Yet, as I have argued in the introduction to this chapter, in order to account for our sense that we are moral agents, who do make moral decisions and engage in moral acts for which we bear moral responsibility, a concept that involves something like moral autonomy is essential. The challenge, therefore, is to develop an adequate account of autonomy; one that, far from denying dependence and development, writes them centrally into the concept. Young and Haste provide hints that part at least of the answer lies in a reconstruction of reason that recognizes that reasoners are not disconnected, dispassionate and disembodied individuals, but must reason from a position in which real flesh-and-blood persons with desires and emotions are inherently situated within a community. In the first section of this chapter, I have delineated a richer multi-aspectual conception - reasonableness - which is capable of describing the way the sorts of people alluded to by Young and Haste actually reason.

In the second section of the chapter, I turned to Kant, briefly outlining his account of reason and the showing how it inevitably leads to the sort of conceptualization of autonomy that Young and Haste have attacked. Subsequently, I surveyed O'Neill's attempt to reinterpret Kant's doctrines in such a way that avoided the charge that Kant considers reason to be a given, that all humans always have, but which, for obscure reasons, children are unable to exercise. This constructivist account I showed ultimately fails, not least because it fails to pay proper attention to the construction of reason in the individual. This alerted us to the need to pay closer attention to developmental questions in the construction of reasonableness and autonomy.

In this passage from Morgan (1996, 249), we see an approach that takes on board the need to recognize attachment and goes part way towards a reconstruction of the notion of autonomy:

Autonomy does not imply extreme individualism or independence from community, but merely that one's relations with other people and one's position in society itself are matters for self-reflection.

This characterization does to some extent recognize connection. It is, however, a connection that seems to arise out of a self already capable of self-reflection. It does not pay any attention to how it is that one *becomes* capable of such self-reflection. It leaves open the question as to whether one is always capable of it, but needs to exercise it in order to become autonomous, or whether such a capability is itself constructed through interdependence. In §3.23, I showed how the work of Piaget offered a number of insights into the development of reason, but that it, too, being based on Kantian foundations, relies too much on the solitary reasoner. In the work of Vygotsky, we found a developmental psychological program that looks to social interaction directly for the roots of the development of reason. Vygotskian learning, involving embodied persons in specifically situated communal dialogue, permeated not only with reasons, arguments and speculations, but also with emotions, commitments and desires, provides for the richness necessary for the development of reasonableness.

This has allowed me to make a first pass at reconceptualizing autonomy. In §3.24, I looked at the notion of thinking for oneself, demonstrating that it does not imply individualistic thinking, but thinking that must both arise, and be judged, in community with others.

In order to complete the task of reconceptualizing autonomy, which will form the final section of the next chapter, it will be necessary to look further into the conditions under which reasonableness can arise. First, I will look at a neo-Kantian program that takes the issue of interdependence seriously, breaking away from the monological reliance on a single agent as the basis for theorizing. This is the work of Habermas on communicative action. This work will allow us to explore the extent to which Kantian insights into autonomy can be rehabilitated. In the course of the exposition of Habermas' project, I will explore the ways in which his account of communicative action can be combined with the developmental perspective identified in this chapter. In particular, I shall concentrate on a modification to his classification of types of action, to take account of pedagogic action: action in which the dialogical partners are not all already competent reasoners. For, as we have discovered of Kant in this chapter, and will discover of Aristotle in Chapter Five, by concentrating on finished humans - adults - Habermas' philosophical project has both remained incomplete and lapsed into error. It is a common failing in the philosophical tradition that theorizing has paid insufficient attention to childhood.

Chapter Four: Pedagogical Action and the Development of Autonomy

During the survey of reasonableness and autonomy in the previous chapter, two points became increasingly obvious. The first is that if we are to make proper sense of what it is to be reasonable, we need to consider reasonableness in a rich contextualized setting in which real thinking and feeling people interact and communicate. The second point is that the best place to look at this is not in the “market place,” where many adults have already become somewhat reasonable and autonomous, but in places where young children are developing reasonableness and autonomy. Given my aim - to look at moral education - it becomes obvious that the place on which to focus my attention now is the classroom, and the activity taking place within the classroom on which to focus is dialogue which genuinely involves the students as well as the teacher.

But first, as I indicated at the end of the last chapter, I will introduce and explicate a neo-Kantian philosophical position that will give me a theoretical basis of analysis of classroom dialogue. This is Habermas’ theory of communicative action.

4.1 Habermas’ Classification of Dialogical Action and the Classroom

The theory of communicative action is Habermas’ attempt to explain how modern human rationality and action arise through the medium of language. Habermas’ theory is, as we shall see both here and in Chapter Six, a neo-Kantian project which takes more account of the place of human interaction and the “linguistic turn” in philosophy. As in Kant’s work, ideals play an important part for Habermas: in particular, the concept of an ideal speech situation plays an important role in the theory of communicative action.

Since this theory has been developing over many years in many works, it is not always easy to come to grips with either its details or its current manifestation. As a consequence, both communicative action and the ideal speech situation have been presented in many different (and not always compatible) guises by different writers seeking to apply Habermas’ insights to particular disciplines.

4.1.1 Critical Theory and Education

Education is one such field. Given the importance of language and communication in education, a number of attempts have been made to co-opt Habermas’ theory for educational ends. Broadly speaking, these attempts to develop a “Critical Theory of Education” fall into two (somewhat overlapping) schools. The first draws on Habermas, usually in conjunction with other members of the Frankfurt School and

various post-modernist writers, to design specific classroom materials and approaches that promote broadly political objectives. This field is often called “Critical Literacy” (see, for example Lankshear and McLaren 1993; Hamston 1995; Pitt, Tugwell et al. 1995; Morgan 1996), and I do not intend to discuss it here.

The second approach has been aimed at characterizing classroom dialogue (or, more broadly, educational action in general) as an instance of communicative action, with the aim of reconceptualizing all such dialogue in Habermasian terms. Often the aim here is also reform: such dialogue is to be critically analysed from the Habermasian point of view and improvements made. Such improvement is less commonly aimed at explicitly political content, as in Critical Literacy, but rather at a reinvigoration of the forms of educational dialogue (admittedly, often with some political implications). Two influential writers in English in this field are Young (1988; 1988; 1990; 1992) and Blake (1995). This discussion has been going on in the continental (largely German) literature for over twenty years, but unfortunately little of this work has been published, or even summarized, in English (see Masschelein 1991; and Peukert 1993 for exceptions). This is the field in which I will operate, with a particular focus on the classroom. I have explained above why I think that the focus of this work ought to be on discussion within the classroom. I do not wish to imply by this that classroom dialogue is all that education is about: clearly there are many other important aspects, but I shall not address them here.

Given the multitude of interpretations of Habermas, I shall begin by sketching in the key relevant features of his theory of communicative action (Habermas 1981; 1990) as I take them. In this somewhat simplified account, I draw on Braaten (1991), Horster (1992), Benhabib (1992), White (1988) and Thompson (1981) in addition to those mentioned in the previous paragraph for assistance in interpreting Habermas. In §4.13, I will outline a plausible application of Habermas’ account to classroom dialogue, while in §4.14, I will critically evaluate its adequacy, leading to the conclusion that certain specific features of educational action provide difficulties for a straight Habermasian analysis of classroom dialogue. Thus I will discuss a further characterization of action - pedagogic action.

4.12. The Key Features of Habermas’ Account

One of the difficulties in getting the key features of Habermas’ account of communicative action straight is the specialized terminology he uses: terms neither have their commonly understood senses nor are used consistently. Here are three examples. Firstly, what I will refer to below as “success-oriented action” seems also to be called, at different places in his works, “purposive-rational,” “means-end,”

“instrumental” or “strategic” action in response to shifts in context. Secondly, he calls his overall theory “a theory of communicative action,” yet one of the specific types of action it describes is also called “communicative action” (I believe this feature alone accounts for a good deal of misunderstanding). Finally, Habermas reserves the relatively common term for focused interactive talk in general - “discourse” - for a very specific purpose, to be explained in §4.123. To avoid confusion, I will use the term “dialogical action” to cover Habermas’ general sense of communicative action, and “dialogue” for focused interactive talk in general.

In the account that follows, I identify four key terms in the theory. The first three refer to specific types of dialogue (or other action) - “success-oriented action,” “communicative action” and “critical discourse” - that can be, and commonly are, used in everyday life. The fourth, the “ideal speech situation,” refers to an explicitly counterfactual situation, unachievable in practice. This contrast between the practical and the in-principle unachievable ideal is a source of some confusion in the interpretation of Habermas, as we shall see. It can also be a target for some important criticisms.

4.121 Success-oriented Action

In success-oriented action, people act as individuals aiming for the success of their own actions, without attempting to coordinate their purposes and ends with others. Habermas (1981, 285, Figure 14) distinguishes strategic (social) from instrumental (non-social) action, though this distinction is not strongly maintained in his subsequent discussion. Strategic action is further subdivided into open and concealed (Figure 18, p 333); he offers no discussion of the open category in the text, presumably because his greatest interest is in a critique of the way concealed strategic action works in modern society.

Given Habermas’ lack of explication, how are we to understand open strategic action? In strategic action, goals are not open to question - merely the means to achieve the goals. Open strategic action presupposes agreement (at least, for the present) concerning the relevant set of transparent goals, facts and values. Open strategic action, then, appears to be open in something like this sense: the intentions of those taking the action are not manipulative (i.e. designed to achieve a goal which needs to remain secret for its success), but rather aimed at achieving a generally accepted goal. This goal has been previously considered (though not necessarily by the participants in the action) and is not, for the present at least, thought worthy of further discussion. The goal is, however, either potentially open to discussion at a level different from that at which the action is taken, or could be opened to discussion at the present level under particular circumstances.

Much of Habermas' discussion of success-oriented action draws on the theory of speech acts developed by Austin and Searle. In that theory, a distinction is drawn between the illocutionary force of a speech act, in which the speaker acts in saying (and so, according to Habermas, the purpose is manifest in what is said), and perlocutionary force of a speech act, in which the speaker brings about actions *by* saying. "Perlocutionary acts constitute a subclass of teleological actions which must be carried out by means of speech acts, under the condition that the actor does not declare or admit to his aims as such" (Habermas 1981, 292). Young (1988) points out that Habermas (1981, 298-303) distinguishes between three types of perlocution. In type 1 perlocutions, the speech act appeals to known normative contexts of legitimate authority, so, for example, a train conductor who shouts "Fares, please" makes a claim for action on the part of passengers. Type 2 perlocutions appeal to known sanctions within the power of the actor, so that a bouncer at a nightclub who asks you to leave implicitly backs up the request with a threat of violence if you don't comply. In type 3 perlocution, the speech act uses deceit with hidden aims, such as when a child says "Look over there," so an adult's attention is diverted so that the brussels sprouts can be fed to the dog. Young claims that Habermas excludes type 1 perlocutions from concealed strategic action, presumably thus characterizing it as open strategic action - an analysis which agrees with that in the previous paragraph.

Often, one actor will be in a position of socially sanctioned authority in order to facilitate the achievement of collective aims. Such open strategic action may, in the interests of efficiency or even sheer achievability, be acceptable and appropriate. For a police officer who is directing traffic around an accident site to discuss the purposes behind these actions with every passing motorist would interfere with achievement of the desired outcome of keeping the roads negotiable. Hence motorists generally trust police to have good reasons for redirecting traffic, even if they cannot see it. Open strategic action is of great interest in education: teachers often seek to find the best way to achieve curricular goals that have been externally set for students.

In concealed strategic action, on the other hand, one participant has goals that are hidden from other participants for manipulative purposes. The most direct example is when a dialogue partner uses words to induce a false understanding about the aims of the conversation in other partners, thus leading them to do as the manipulator desires. A police officer may lead motorists to believe that traffic redirection is due to an accident, when the aim is rather to keep witnesses away from an illegal police operation, or to allow space for disarming of a terrorist bomb without creating

widespread panic. As the second outcome of this example shows, even concealed strategic action does not always lead to consequences that are not in the interests of the deceived party, though the first shows that it certainly can.

Habermas does not claim that all success-oriented action is to be avoided: on the contrary, he claims that it plays an important role in many areas (e.g. scientific research, achievement of socially mandated aims). It is the use of success-oriented action in situations where communicative action or critical discourse ought, in the interests of all concerned, to be used to which he objects: “all three rationality complexes... have not found a balanced institutional embodiment in... modern societies” (1981, 273). For example, Australian social workers in the past attempted to create a better future for mixed race Aborigine/white children by taking them from their Aboriginal mothers and placing them with white foster families. This was a bureaucratically decided solution to the problem of improving the life prospects of certain people, not one negotiated in a political arena that included effective involvement of all affected. He refers to this extension of success-oriented action to situations in which it is inappropriate as the “colonization of the life-world.”

Success-oriented action (open or concealed) is, for reasons of practicality, widespread. Nevertheless, the aims of such actions, even when they are taken in what is perceived to be the best interests of all concerned, can be opened to question. This questioning can be initiated by either the participants themselves, or by those beyond the situation, such as general citizens. If such questioning is never allowed, then success-oriented action often leads to oppression. When questions that go beyond the means to achieve ends are raised, dialogue needs to move to another level, that of communicative action.

4.122 Communicative Action

According to Habermas’ account of communicative action, the parties in such a conversation seek to exchange information, convey experiences and/or clarify meaning through hermeneutical inquiry, coming to a common understanding against a background of assumed facts (truth claims) and values (claims of normative rightness). Communicative action thus does not merely employ means-ends rationality, but aims at reaching intersubjective understanding. Communicative action allows, *inter alia*, for the goals of success-oriented action to be called into question, and for the establishment of common goals.

In policing, for example, someone might ask whether police should be spending more time walking the beat, rather than directing traffic. While such a discussion takes place against a background of shared assumptions about the value of both a

police presence on the streets and of traffic control, and of facts concerning street policing and traffic direction, then the dialogue is conducted within the domain of communicative action.

Similarly, if a number of participants in a dialogue are jointly trying to make sense of each other's positions on some question, they enter into a hermeneutical circle of inquiry, where each attempts to "merge horizons" with the others. Such an inquiry creates a shared intersubjective grasp of reality and society which can both result in the incorporation of the participants into the community (socialization), and lead to some flexibility in communal understanding (change and innovation).

In communicative action, it is assumed that participants meet four conditions: they (1) use language correctly; (2) are sincere; (3) speak the truth; and (4) conform to accepted social norms. Speakers may, however, be mistaken, and it can also happen that, due to ideological or hegemonic factors in society, all participants in a dialogue may accept falsehoods. Any such claim can be called into question. It is Habermas' claim that the attempt to come to agreement about the answer to such questions is a different form of dialogical action: critical discourse.

4.123 Critical Discourse

Critical discourse is the attempt to reach consensus about contentious claims for normative correctness or truth:

Under the heading 'discourse,' I introduced the form of communication characterized by argumentation in which validity claims that have become problematic are made the subject of discussion and are examined relative to their legitimacy. In order to entertain a discourse, we must in a sense step out of behavioral and experiential contexts; here we exchange not information, but rather arguments that serve to establish (or reject) problematic validity claims. (Habermas 1971, 130-1; translated and cited by Horster 1992, 31)

Each dialogue proceeds in the context of implicit background claims of normative appropriateness and truth ("behavioural and experiential" in the quote above). The dialogue could not get started without this, and most such claims are accepted by the participants in the dialogue. Nevertheless, these claims can be called into question - not all at once, but piecemeal. When a truth or a normative claim is questioned, the participants can try to justify or redeem (respectively) the validity of the claim: that is, to reach, through argumentation not directly aimed at action, an understanding about the truth or falsity, appropriateness or inappropriateness of the claim.

Theoretical discourse is critical discourse centred on truth claims. The aim of theoretical discourse is to justify the validity of the truth claim by reaching consensus. Habermas offers a model for such justification, drawing on the work of

Toulmin (Habermas 1981, 24ff; Thompson 1981; Horster 1992, 36-8). When an assertion is questioned by the dialogical partner, *data* (for events: causes; for actions: motives) must be offered, supported by a *warrant* in the form of empirical uniformities or hypothetical laws, that have the *backing* of observations, assessments or other empirical evidence. It is in negotiation of these proffered grounds for belief that the truth claim is established or rejected. This judgement is in itself provisional: when the participants in the theoretical discourse reach consensus as to the validity of the claim it recedes into the background, though it may be challenged again in the future.

The attempt to redeem the validity of claims about normative appropriateness (that is, practical discourse) follows a similar pattern. Once again, the participants in the discourse strive for consensus when a normative command or prohibition is questioned. The data in this case take the form of the grounds for the judgement, which must be warranted by behavioural or evaluative norms or principles, themselves backed by reference to needs, values, consequences and ramifications. Again, consensus is provisional.

Critical discourse requires one to distance oneself from the implicit background that underpins our situation in the world (Heidegger 1962). It requires a degree of reflexivity in order to be able to problematize the taken-for-granted. (It is worth noting that Habermas' insistence on dialogue rather than monologue springs in part from the need for other points-of-view to facilitate distancing and problematizing). And while the *raising* of contested validity claims merely requires a capacity to express interests, the *redeeming* of these claims in critical discourse also requires a degree of sophistication in reasoning, so as to be able to adduce data, advance warrants and support them with backing. All these are relatively high level capacities which draw upon a considerable degree of intellectual prowess.

Habermas claims that truth (in theoretical matters), correctness (in practical matters) or even rationality itself just is what would be agreed upon by an ideal speech community in the ideal speech situation and, moreover, that the legitimacy of discourse itself depends on the presupposition of the ideal speech situation.

4.124 Ideal Speech Situation

Unlike the three types of dialogue we have been considering, the ideal speech situation cannot be entered; it is counterfactual. It is Habermas' description of:

the general pragmatic presuppositions of argumentation [i.e. critical discourse]... [an attempt] to reconstruct the general symmetry conditions that every competent speaker must presuppose are sufficiently satisfied insofar as he intends to enter into argumentation at all. Participants in argumentation have to presuppose in general

that the structure of their communication, by virtue of features that can be described in purely formal terms, excludes all force... except the force of the better argument (and thus that it also excludes... all motives except that of a cooperative search for the truth). From this perspective, argumentation can be conceived as a *reflective continuation, with different means, of action oriented to reaching understanding* [i.e. of communicative action] (1981, 25, italics in original).

First, there are conditions of participation. Benhabib (1992) captures these in two universal pragmatic moral rules: universal moral respect (everyone has the right to be included in the discourse) and egalitarian reciprocity (there is an equal right to make assertions, ask questions of others, introduce new subject matter, call the validity of claims into question and so on). For an alternative formulation, see White (1988, 56). Second, four conditions must be met:

1. each person speaks intelligibly - that is, their assertions accord with the rules of language;
2. each person taking part in the conversation is sincere;
3. assertions made about the existence of certain states of affairs are true;
4. assertions made conform to socially expected standards appropriate to the interpersonal context in which they are made.

It is important to note that the ideal speech situation does not spell out the necessary conditions for critical discourse, for several reasons. Firstly, the ideal speech situation is counterfactual whereas critical discourse frequently occurs in reality. For example, the former requires the presence of all possible speakers, whereas critical discourse may involve only a few speakers. Secondly, the ideal speech situation requires the four conditions listed above to be met, while critical discourse occurs precisely because at least one participant has explicitly questioned whether one of them *is* being met. Nevertheless, although critical discourse explicitly fails to meet the conditions of the ideal speech situation in these (and other) ways, it still requires its anticipation.

Before expanding what is meant by such an anticipation, it is worth noting that the emphasis on finding a single consensus on matters of truth and validity can be questioned (Benhabib, 1992). There is a lively discussion in the literature as to whether Habermas' ideas can be adequately modified to take account of diversity (see, for example, many of the contributions to Meehan 1995), and I shall return to this matter in §6.2322.

The concept of an ideal speech situation needs to be treated carefully. Its counterfactual nature clearly points to the Kantian roots of Habermas' account, and it is well to remember how Kant treated such transcendental ideas. He did not, of course, discuss the particular issue of the ideal speech situation, but his comments about transcendental ideas in general will help us to see Habermas' drift. Here is Kant's own warning about their use:

Transcendental ideas have their own good, proper and therefore *immanent* use, although, when their meaning is misunderstood, and they are taken for concepts of real things, they become transcendent in their application and for that very reason can be delusive. (Kant 1965, 532, A643/B671)

Kant cautions us that transcendent ideas, ideas that arise from consideration of the logical conditions which must hold in order to make our experiences possible, should not be mistaken for a description of a real object. So, if the ideal speech situation is not to be taken for the concept of a real dialogue, then how is it to be interpreted? Again, according to Kant:

[Transcendental ideas] have an excellent, and indeed indispensably necessary, regulative employment, namely, that of directing the understanding towards a certain goal upon which the routes marked out by all its rules converge, as upon their point of intersection. This point is indeed a mere idea, a *focus imaginarius*, from which, since it lies quite outside the bounds of possible experience, the concepts of the understanding do not in reality proceed... (533, A644/B672)

Transcendental ideas, Kant argues, have an imaginary, counterfactual focus, not an achievable one, but that they can assist us by regulating what we do set out to achieve, as guiding epistemic principles. The ideal speech situation, then, is to "direct our understanding" rather than to be taken as a real achievable model for dialogue. Habermas (1992, 144) clearly has this in mind:

The idealizing presuppositions of communicative action must not be hypostatized into the ideal of a future condition in which a definitive understanding has been reached.

In the following passage (where the principles in question are those of the homogeneity, variety and affinity of the appearances), Kant provides three ways in which the ideal speech situation might regulate discourse.

The remarkable feature of these principles... is that they seem to be transcendental, and that they contain mere ideas for the guidance of the empirical employment of reason - ideas which reason follows only as it were asymptotically, *i.e.* ever more closely without actually reaching them - they yet possess, as synthetic *a priori* propositions, objective but indeterminate validity, and serve as rules for possible experience. They can also be employed with great advantage in the elaboration of experience, as heuristic principles. (545, A663/B691)

I will elucidate the meaning of each of the three aspects Kant mentions by seeing how each applies to Habermas' ideal speech situation. Kant's first way concerns

“guidance of the empirical employment.” Applied to dialogue, this implies that each time we enter into speech, we not only should try to approach the ideal speech situation as closely as possible, but also need to presuppose (in the absence of counter evidence) that others involved are doing so: that is, we must enter dialogue in trust. Each dialogue is, if not deliberately distorted, a striving towards perfection, but is limited by contingent factors so that it can never “actually reach” that goal. Given the contingent facts about our world (e.g. our mortality, our limited reason), then this approach will not in fact be very close. I shall call this way of considering the role of a regulating ideal the “existential aspect.”

When we consider the conditions of the ideal speech situation as “synthetic *a priori* propositions... [which] serve as rules for possible experience,” then we begin to see their role as preconditions for speech. In Kant, the best parallel is the role of freedom in moral action. Kant asserts that his transcendental claim - that the will is free - is unprovable, since such a will could only be *noumenal*, and we have access only to the *phenomenal*. Yet we must act under the idea of freedom, for if we do not, then we can only think of ourselves as determined, and we can make no sense of how it is that we act voluntarily and responsibly in the world. Habermas makes a parallel claim: the assumption of the ideal speech situation is what makes possible all real dialogues. The mere engagement in argumentation, the attempt to contest and determine validity or truth, presupposes that, under ideal conditions, a consensus could be reached (see §6.2322 for further discussion). If not, then the rationale of the engagement disappears. A skeptic who engages in an argument, while professing to deny the presupposition that it can possibly lead to agreement, is caught in a performative contradiction (Habermas 1990, 88-92, see also §6.1). We must consider dialogues to be aimed at cumulatively achieving an overarching ideal speech situation through the continual refinement of previous dialogues. If we do not, then the point of a single dialogue, even one far removed from the achievement of an ideal speech situation, disappears. This I’ll call the “teleological aspect.”

Finally, Kant refers to the role of transcendental ideas as “heuristic principles” for the “elaboration of experience.” For Habermas, the ideal speech situation provides us with a checklist of desirables (rather than an algorithm or set of rules to be followed exactly) as a way of keeping tabs on, and making sense of, our attempts to engage in discourse (Habermas 1990, 92). Without presupposing the conditions of the ideal speech situation, we would have difficulty in evaluating (either procedurally, as Habermas argues, or indeed in qualitative research terms) those real speech situations in which we do constantly find ourselves. They provide a regulating ideal against which to measure our attempts, in something of the same

way as the counterfactual “ideal second,” unachievable in practice, provides a unit against which actual times can be measured, and deviations corrected. This is the “practical aspect.”

Each of these ways of looking at the role of the ideal speech situation enters into Habermas’ characterization in different contexts, and they raise serious problems for those who take the ideal speech situation as a set of rules for the conduct of classroom dialogue. Firstly, in the Habermasian account, not all classroom dialogue is even subject to the conditions of the ideal speech situation since, as Blake (1995, 356) points out, they are only relevant to critical discourse. Once we realize that Habermas is using the word “argumentation” interchangeably with “critical discourse,” this is obvious from the quotation with which this section commenced. But even in the case of critical discourse, caution is needed. Kant comments: “When merely regulative principles are treated as constitutive, and are therefore employed as objective principles, they may come into conflict with one another” (547, A666/B694). He is warning us that transcendental ideal conditions are not rules for constructing real objects, and that if treated as such they may be contradictory. Thus, for example, if we insist that a classroom dialogue allow any participant to introduce any material at any time, this will ensure that the dialogue cannot come to consensus within any feasible time allowed. Blake (1995) also makes this point: “Habermas’ prescription for an ideal speech situation can seem unworthy of even partial emulation. For what seems enjoined is an almost anarchistic¹ conduct of discussion... unstructured speech situations can kill rationality.”

So, it is easy to make the case that the ideal speech situation should not be something we strive to implement in reality in every respect, for those respects may conflict. Some of the respects are impossible (such as allowing all possible speakers into every conversation, or always knowing the truth so as always to speak it) and others, as we have seen, are unwise or counter-productive even if the conversation is limited to the physically present participants. Yet the existential aspect does give us some guidance: when we must, for reasons of contingent limitations, move away from Kant’s “asymptote” by imposing constraints that run counter to the ideal conditions, we need to be wary of the ways in which these constraints may contaminate our deliberations. Habermas (1981, 25) says that “the structures of an ideal speech situation immunize against repression and inequality in a special way.” One example: if we exclude some from participating in a particular dialogue, or ignore certain comments made by some participants, it may make the dialogue more focused and easier to run, but it may also undermine the conclusions reached. If we

¹ Perhaps it would have been better if Blake had chosen the word “chaotic” rather than “anarchistic” here. Further, while unstructured discussions can damage rationality, they need not.

analyse the dialogue under the practical aspect, by noting that universal moral respect, or egalitarian reciprocity has been ignored, we can help to make these distortions clear.

Nor can we consider that all the problems with the ideal speech situation in specific dialogues can be resolved by looking to the “mega-discourse” implied by the teleological aspect. While access to all, including those separated by time and place, can be alleviated by “institutionalised discourse” (Blake 1995)², it cannot be completely overcome: for example, earlier participants have no opportunity to challenge latter ones (particularly if they have already died), and even widely spread institutions are subject to hegemonic controls.

In the context of this discussion, where I am trying to make sense of the practice of dialogue in classrooms, the most useful aspects are the existential and practical aspects. For Habermas, when we sincerely enter into discourse, we must existentially act *as if it were the case* that the conditions of the ideal speech situation were being met, even though we know they cannot be. We are expressing the hope that our claims and assertions, and our defences of them, are such that they would be capable of achieving universal assent, even while recognizing that this may be a practical impossibility. It is this that enables us to make sense of the enterprise and allows us to accept the outcome of the discourse:

[In each discourse we] mutually assume an ideal speech situation. The ideal speech situation is characterized such that each consensus that can be achieved under its conditions is valid as a true consensus per se. The anticipation of the ideal speech situation is a guarantee that we may associate the claim to a true consensus with a consensus that has actually been attained. (Habermas 1971, 136; translated and cited by Horster 1992, 34)

We act in good faith, attempting to the best of our ability to make true and normatively valid claims, and make the assumption that all other participants are acting likewise. Even when we (or others) enter into discourse insincerely (and thus engage in concealed strategic action), it is the underlying presupposition of the ideal speech situation (and our adherence to it in many ways, even in the midst of dissembling) which allows the deception to work. For we enter any dialogue, even a strategic or communicative one, under a “speech-act immanent obligation”; that is, an expectation that we are able, if required, to justify or redeem any validity claims that underlie our statements. Since such justification or redemption would have to be

² It is worth noting that Blake's main aim in his paper was not to analyse classroom dialogue in Habermasian terms. Rather, after pointing out difficulties in implementing Habermas' ideal speech situation in a classroom, he made the claim that it is to institutionalized, rather than specific classroom, dialogue that Habermas' work is most applicable. However, even institutionalized dialogue cannot meet the standards of the ideal speech situation, and we should not expect it to.

carried out in critical discourse, there is a sense in which the ideal speech situation is anticipated in all dialogical action.

But although we must always presuppose the ideal speech situation in this way, in order to enter into dialogue, we ought also to be aware through the practical aspect of the ways in which the actual speech situation falls short of the ideal speech situation (since it always must), and to make judgements as to whether the actual speech situation is adequate to the purposes of the dialogue. If it is not, then we have the possibility of moving it to another level of dialogical interaction. Thus strategic action can become communicative, and communicative action can shift to critical discourse.

Habermas claims that it is only through the possibility of such an ascent through the modes of dialogue, ultimately underpinned by the presupposition of the ideal speech situation, that each mode can be legitimated. That is, strategic and communicative dialogue play essential parts in our coordination of social activity, but each on its own can, and does, lead to distortions and oppression if unchecked by the possibility of critique. Critique itself relies for its justification on the teleological aspect of the ideal speech situation; that is, on the underlying presupposition that consensus is the (counterfactual) *telos* or goal of dialogue. Although such a strong presupposition of consensus has been questioned (see a number of papers in Meehan 1995; especially Warnke 1995), Benhabib (1992) and Braaten (1995), amongst others, have argued that Habermas' account can be adequately modified to incorporate an additional *telos* of mutual recognition and understanding. I shall return to this matter in Chapter Six.

4.13 Types of Dialogue in Classrooms

Equipped with this understanding of Habermas' theory of communicative action, it is time to turn attention to the implications for the classroom. In this section, I will explore whether Habermas' theoretical framework can help us to make sense of the actual use of dialogue within the classroom, and in the next, whether it needs modification. Once we have an adequate theory of dialogue in classroom situations, it can play an important role in the analysis and critique of the use of dialogue, leading to the development of sound, theoretically based improvements in practice.

At the base of the problem to be addressed in this exploration lies what is known in the German and other continental literature as the "pedagogical paradox." Perhaps the earliest formulation of this paradox lies in Kant's question: "How shall I cultivate freedom through force?" (*Über Pädagogik* p. 453, cited in Kivelä 1998; see also Kivelä, Peltonen et al. 1998). In other words, how can teachers justify unilaterally imposing many activities and restrictions on children if one of the important aims of

education is to produce citizens who can think for themselves: that is, reasonable and autonomous citizens?

The question of whether Habermas' categories will turn out to be useful for this purpose must wait until we have seen how well they can be applied to classroom dialogue. There are, however, several good reasons for supposing that they might. Firstly, they enable us to focus on the degree to which purposes are disclosed and negotiated within dialogues. Secondly and relatedly, they explicitly distinguish actions which are imposed on other dialogical partners from actions which are agreed between them, as well as further identifying actions involving the exploration into the validity of the presuppositions that underlie these actions. Thus questions of manipulation and freedom can be directly addressed.

§4.12 should have made it clearer why Habermas' work is of especial interest to me. Education is a social institution, within which the teacher has a role to play. Communication with the students is one essential part of this role. The teacher, however, is not just another partner in communication: part of the social role is to structure the educational environment to achieve socially mandated aims, high amongst them induction into the communication community. This social responsibility of teachers entails that teachers are mandated to wield the power necessary to imposing activities on the students. The exercise of this power is sanctioned provided it is in certain specific interests of the students: I shall explore this issue of power further in §4.14. Thus the teacher may act strategically, seeking to achieve curricular aims without jointly negotiating them with the children.

Yet this is a curious sort of strategic action. The aims of the teacher are not personal and selfish. They are only achievable if the students become (amongst other things) proficient members of the adult communication community. This aim, though not mutually negotiated with students (in fact, being, at least in modern societies, one of the non-negotiable foundations of education as an institution), is certainly open to question at a number of levels, to be discussed later. If we focus on the teacher's role, Habermas' categories provide a means of justifying and legitimating the action taken in terms of the purposes of classroom dialogical interaction. His categories of dialogical action identify three levels: action where the goals are unilaterally implemented; action where the goals are jointly constructed; and finally, action where the presuppositions that underlie the goals can be opened to question in pursuit of legitimation of those goals. This promises to be useful in providing an ethical underpinning to the teacher's actions in the endeavour of education, especially education in ethics, in ways which will be explored in §4.14.

Furthermore, if we focus on the students, Habermas' categories allow us to distinguish between dialogue which instantiates goal-directed actions foisted upon them, in which they have no say, dialogue which involves them in (at least a degree of) joint consideration of the goals and finally the critical examination of norms and facts underpinning goal setting. At the core of the process of education is the endeavour, through a supported process of development, to move from a situation in which young and unskilled children need considerable outside direction, to a state in which they have acquired the ability to participate fully on their own account; that is, education attempts to move students towards autonomy.

It is important to keep in mind the scope of my inquiry. As specified in §4.12, the term "dialogue" is being used in a restricted sense, to refer to focused interactive talk. Thus, I will not consider unfocused chat, one-way instruction and the like (nor will I attempt to draw a clearer distinction between these and focused interactive talk). Indeed, for the purposes of exposition, I will narrow the scope (and hence the sense of "dialogue") slightly further, to those episodes of classroom dialogue that teachers commonly refer to as discussion: episodes where children are both allowed and encouraged to talk interactively with the teacher for directly educational purposes. I am not claiming that Habermas' categories are inapplicable to all these other types of classroom talk: I shall leave this possibility unexplored. Rather, I am restricting my inquiry to the sort of talk in which I am most interested, for its undervalued educational potential (Dillon 1994).

We saw in §3.324 that any assertion that dialogue in classrooms ought to meet the conditions of the ideal speech situation must be misguided, since no real dialogue can ever meet those conditions. Nevertheless, for Habermas, the counterfactual presuppositions of the ideal speech situation underlie (even though they don't constitute) *all* dialogue, including classroom dialogue. *No* classroom (or other) dialogue can actually achieve the ideal speech situation, but all such talk can and does achieve one of the three levels of dialogical action, or possibly (as I shall explore) a mixture of several of them. We can use the two universal pragmatic rules (universal moral respect, egalitarian reciprocity) and four conditions (intelligibility, sincerity, truth, normative correctness) of the ideal speech situation to guide us in our critique of actual classroom dialogue, but we cannot demand that it meet them.

I will now turn to the task of considering actual dialogues within schools with respect to the three levels of dialogical action. It is not difficult to see that much classroom dialogue comprises of success-oriented action. The teacher is in a position of power within the classroom and controls talk for pedagogic purposes. Much empirical research into classroom dialogical interaction reveals that students'

responses to teachers are often of the form of “guess what the teacher wants us to say.” Dillon (1994) calls such questioning *recitation*. He notes that, although teachers commonly call this form of classroom interaction “discussion,” it consists of cycles of closed, shallow teacher questions, brief student responses and teacher evaluations. Sinclair and Coulthard (1975) refer to the same pattern as an IRF cycle (initiation, response, feedback) and present evidence that most classroom interaction fits into this pattern. Lemke (1990) also finds the same pattern dominating the science classrooms he studies, although he calls it Triadic Dialogue. Of especial interest in this context is the work of Young (1992, Chapter 8), who also finds a preponderance of IRF interaction in classrooms, and explicitly discusses it in terms of strategic action. I shall return to his analysis in §4.14.

In such classroom dialogue, the teacher has a clear aim in mind and controls the dialogue tightly in order to achieve it. Although less commonly used nowadays, straight lecturing and dictation of notes are further pedagogic methods designed to meet certain teacher aims, as are such skills reinforcement techniques as the chanting of tables. The point to be made is this: although all these teaching techniques can be characterized in Habermasian terms as open strategic (success-oriented) action, each of them has its own justification in appropriate circumstances. Recitation can be an effective way to check students’ understanding. Lecturing is a highly efficient way to cover content when the students are motivated and properly skilled. There are very good reasons for being able to recall “number facts” instantly. It would make little sense to call for the avoidance of all success-oriented action in classrooms, just as it would make little sense to banish success-oriented action from wider society. As we saw in §4.121, Habermas makes this point, although it often gets lost in his critique of the inappropriate extension of success-oriented action into the life-world.

However, a good empirical case can be made for the claim that such success-oriented action leads to shallow, poorly integrated learning, and that it takes more than its fair share of classroom time. This is effectively the point made by Dillon, Sinclair and Coulthard, Lemke, and Young (op. cit.). They, and many others, call for more interaction that allows students to express their present understandings, check them against those of the other students, the teacher and the world and rebuild them if necessary. Such interaction seems to fit well into Habermas’ category of communicative action: indeed, if the students in the course of such inquiry problematize truth or validity claims, they enter critical discourse as well. Often referred to as social constructivist learning, such techniques have attracted considerable support for their educational value (see, for example, Wells, Chang et al. 1990).

Such episodes have been identified in classroom dialogue. Lemke (1990), for example, found dialogical types he labelled True Dialogue (where the teacher asks open questions and the students explore the possible answers) and Cross-Discussion (dialogue carried out directly between students, with the teacher either moderating or participating as a [quasi]-equal). Nevertheless, he points out that, of all the dialogue types he identified, these are the two rarest, and that teachers, who afford them little value, commonly bring them to an early close. Both Young (1992) and Dillon (1994) provide detailed analyses of how discussions that embody communicative action and even critical discourse can be created in the classroom, providing transcripts of real classroom discussions that achieve these aims.

4.14 A Critique of Habermas from the Classroom Perspective

This outline of an application of Habermas' ideas to classroom dialogue skates over a number of difficulties. If the theory is to be used in analysis and critique of classroom dialogue, these must be addressed. Some of the difficulties may be generalizable to wider applications of his work, but I shall confine myself here to a critique of Habermas' theory from the classroom perspective.

Firstly, it is immediately obvious that any attempt to attach to a complete classroom dialogue a label stating that it falls under the category of a single one of the three types of action is bound to be too coarse, at this level of analysis. We have seen that if children are, for example, engaged in the hermeneutical inquiry into shared meaning of communicative action, it is likely that disputed claims about what is actually the case, or what is right to do, will arise. Thus communicative action will shade into critical discourse, and just as likely return to that level. Similar shifts from strategic to communicative action and vice versa can also be expected.

Secondly, without intervention and cuing from the teacher, much classroom discussion is likely to be relatively unfocused and children are unlikely to learn the skills and abilities (including the ability to remain focused) necessary to take part in each type of dialogical action. These capacities are the ones that I labelled reasonableness in §3.1, though, of course, reasonableness can be used in many other settings than just dialogue. Remember the stringent requirements for discourse, either theoretical or practical. Children are in the process of learning these complex dialogical competencies, and it is the task of the teacher to assist them. Of course, not all teachers see it this way: if the control they exert on the classroom ensures that the dialogue seldom emerges from success-oriented action, the children will be impoverished in their abilities to engage in other forms of dialogical action (Lipman 1991).

This second point leads to a wider observation. According to Habermas, success-oriented action is appropriate when it is utilized in the pursuit of socially mandated aims. We might consider that educating children is one such aim, and that the profession of teaching is dedicated to achieving that aim. Thus we could claim that success-oriented action is perfectly acceptable in the teacher's role. Further, the teacher is in a position of responsibility to bring about the consummation of that aim, a position that places the teacher in an unequal power relationship with the students in the school. The analysis of the previous paragraph shows that this is necessary for education to take place (I will further consider the question of power below). Some traditional, conservative educational commentators argue that this mandates transmission-style education and teacher imposition of curriculum; education as success-oriented action. Young (1992, 54-5) cites R. S. Peters, R. F. Dearden, R. Spaemann (in explicit reaction to German critical theory of education) and even John Stuart Mill as among those who argue for a mastery of the bodies of knowledge before engaging children in the joint establishment of goals and understandings with others.

The situation is, of course, more complex in several ways. If education had a single, unitary aim, there might be some prospect of transforming society-wide agreement that is a Good Thing into a purely success-oriented role for teachers. But "education" is itself an essentially contested term, and it is analysable into many different sub-units. Two commentators who agree on the vital importance of education may be advocating quite different things. Even to understand the point of view of the other requires communicative action. Given a wide divergence in views on what the facts and the values are that underlie education, entry into critical discourse on the matter is inevitable, at least at the level of those whose task it is to decide on what to deliver as education.

Yet, in a democracy, those who decide on educational aims are, at least in theory, the public. Since this public is itself produced (in part, at least) through the education system, then there is a need to equip all children with the competencies to engage in communicative action and critical discourse. Here is where the apparently clear distinction between the *prima facie* success orientation of education and the need to engage in communicative action and critical discourse begins to collapse. For there is considerable weight to the argument that these latter two modes of dialogical action can only be learned (or at the very least, can be learned best) through engagement in them (Lipman 1991; Sprod 1994b, §3.23). At the same time, if they are being learned, then the children engaged in them are not yet competent, and they need guidance and scaffolding from the teacher (Bruner 1986).

These considerations suggest that, while the teacher is joining the children in (say) critical discourse, they ought simultaneously to engage in the success-oriented action necessary to teaching children how to do it better. This is the view I shall be defending in the rest of this section. This is because the teacher is deliberately structuring the situation in certain ways with the aim - not exactly hidden but usually not made explicit either - of equipping the children to engage more competently in critical discourse. The teacher sets up the conversation, decides on the text or broad subject matter used to start it (even if not the specific questions to be addressed), makes certain interjections into it (such as asking questions, making comments, nodding, smiling, looking puzzled) and requires certain behaviours (e.g. listening to the speaker, respect for others - even being present in the classroom). Each of these is done with a dual purpose: to advance the inquiry; and to improve the ability of children to participate in such inquiry. Some of the interventions are clearly deceitful, such as the use of Socratic irony (the pretence not to know the answer to, or have opinions on, a question), yet this deceit clearly has an educational justification (not to stifle or pre-empt student inquiry). Others are just as clearly coercive, backed by the ever-present possibility of sanctions and the institutional power of the school.

The account I am developing here bears quite a few resemblances to that advanced by Young (1992), in a book that has the advantage that it deals not only with theoretical arguments, but also detailed analysis of real classroom dialogue. In one instance, he considers a discussion transcript that he characterizes as discourse, and comments:

The agent roles of this form of collaboration are roles of joint rational responsibility, although the teacher's role remains one of 'leadership' in the sense of taking a special, continuing and methodological responsibility... for critique of the quality of pupil reasoning.... There is still talk asymmetry in this classroom, but it is a different asymmetry from that in the dominant [IRF, strategic] classroom type. This asymmetry is complementary, since the rights of pupils as rational interlocutors are preserved, while the teacher's superior knowledge and rational skills are still able to be employed on the pupils' behalf in the fostering of the inquiry of the class. Complementary asymmetry of this kind is... what justifies the presence of a teacher. (ibid., 117-8, square brackets my clarification)

Young is at pains to emphasize (and I agree with him) that this is only one of the approaches to teaching that teachers are justified in using. He characterizes strategic action in education (ibid., 48) as the gaining of students' acceptance of claims without their rational assent, solely due to the teacher's authority. This, he says, can be justified by, amongst other things, the circumstances of the classroom, the particular aims being addressed and the curricular and institutional constraints on the teacher (ibid., 108; see also my discussion of indoctrination in §9.4).

Yet, if the teacher is going to move beyond pure strategic action at all, this peculiar mix of success-oriented action with communicative action and critical discourse seems to be inescapable. If teachers abdicate the overall control of the educative situation by becoming mere participants in communicative action, then they are no longer teachers. Totally unstructured interactions between children can of course be educational in a wider sense, but they are not *teaching* situations. To empower children does not mean handing over power to them willy-nilly, allowing them to decide everything that goes on in the classroom, for this would lead in many cases to its misapplication. Rather, empowering means developing in people the capacities to be able to act powerfully on their own account (i.e. autonomously), and this requires positive and purposeful interventions.

Such accounts require an analysis of power, coercion and deceit which legitimates their use in pedagogical situations, whilst highlighting the undoubted dangers that accrue. Power is often thought to be inherently repressive, imposed in top-down fashion by the powerful on the powerless. Yet Foucault (1991) reminds us that:

the notion of repression is quite inadequate for capturing what is precisely the productive aspect of power.... What makes power hold good, what makes it accepted, is simply the fact that it doesn't only weigh on us as a force that says no, but that it traverses and produces things, it induces pleasure, forms knowledge, produces discourse. It needs to be considered as a productive network which runs through the whole social body, much more than as a negative instance whose function is repression. (60-61)

Since the negative guise of power is repression, then its productive use is its positive guise (though we need to note that, for Foucault, not all knowledge that power produces is good for those involved). There are two important points in this passage: power in its positive guise is essential to the production of knowledge; and power is distributed throughout social interaction. If schools and classrooms are merely sites of the application of repressive power (and some of the worst of them certainly approach this), then the educative endeavour breaks down. Secondly, Foucault reminds us that even in such situations, power is not merely concentrated in the teacher's hands, but is distributed, albeit unevenly, throughout the classroom, to appear in every interaction. In repressive schools, for example, the efforts of the teacher can be thwarted by the students' power to disrupt or withdraw.

Wartenberg (1990), in his analysis of the nature of social power, also identifies a positive guise for power, with explicit reference to the classroom. He says, "in so far as one treats other human beings in a way that encourages them to seek to develop their own potential, one can engage in transformative power relationships with them" (214).

Good education requires that students (as they become increasingly able to do so) voluntarily accept the teacher's coercive power, and trust the teacher to use it wisely.³ The teacher's power is exercised legitimately as long as it is used for Foucault's productive purposes (creating student knowledge), *and* it contributes to the students' interests; that is, it is Wartenberg's transformative power. The students' interests will initially be decided by others on their behalf, but increasingly these decisions will take account of the students' own input. Central amongst these interests is the induction of the students into the communicative community, without which they are unable to engage in this critique of what their interests are, or of how these interests can be coordinated with those of others.

This point - that students are initially incapable to a large degree of either deciding or representing their interests - pinpoints one of the dangers in teacher power. As Wartenberg observes, "The crucial thing to recognize is that the presence of power relationships causes human beings to make choices that determine the sorts of skills and abilities they will come to have" (160). In doing this, such power will influence the interests students come to have, for having these skills and abilities are important in deciding and revising one's interests. We might well add that teachers directly influence the things that students become interested in, and how they frame their interests. What all this amounts to is the possibility that a teacher's transformative power can be misused, a possibility usually marked by the term "indoctrination." I shall return to this danger briefly towards the end of §4.141, and treat it in more depth in §9.41.

Hoy (1986, 134) expands on Foucault's account of the point of action of power: "As a complex strategy spread throughout the social system in a capillary fashion, power is never manifested globally, but only at local points as 'micro-powers'." If we accept this, as I think we must, an analysis of teacher power (and its relation to student power) needs to be carried out at the micro-level, a task for classroom dialogue analysis rather than theoretical analysis. Nevertheless, some general observations can be made. To choose an example central to the concerns of this chapter, when the class is engaged in discussion, I will argue that the teacher's power is more likely to be legitimately exercised if it is confined in two ways.

Firstly, the teacher's power ought to be exerted more in process matters (such as encouraging the giving of reasons, the recognition of assumptions, the drawing of valid conclusions) than in tightly controlling content (the actual matters that the discussion shifts to consider). The requirements of reasonableness in communicative

³ For a fuller account of educative coercion and trust, albeit in the slightly different context of tertiary education, see Garfield(1997).

settings are not tied essentially to any particular subject matter (although we have seen in §3.124 that they are most certainly context sensitive). They are, however, heavily dependent on certain processes, as was seen in §4.123, where the onerous criteria for discourse were surveyed. Thus, the teacher's control is pedagogically legitimate when it leads the students to internalize these processes, Vygotskian-style (§3.23), for reasonableness is not only in the interests of the students; it is central to the students' ability to form and critique their interests at all. Of course, not all processes contribute to reasonableness, so that teaching students to stereotype, or to use *ad hominem* arguments would not be legitimate.

Some control over the general content of discussion is necessary for a variety of pedagogic reasons: to achieve certain curricular aims; to keep the discussion focussed sufficiently that variety and depth of communication can be achieved; to take seriously comments previously input into the discussion and thus value that contributor; to allow many to address a particular issue before moving on; or (more generally) to build social and functional competence in a range of areas essential to the ability to act meaningfully in society, amongst many others. In some of these, it can be seen that a sharp process/content distinction is not always tenable.

Nevertheless, excessive content control runs the real risk of collapsing apparently communicative action to pure strategic action. Communicative action needs to be aimed at mutual understanding. As characterized by the practical aspect of the ideal speech situation, it can be mutual only when all present have a chance to introduce material into the discussion. Practical considerations (such as those mentioned at the start of this paragraph) limit the extent to which these freedoms can be allowed in real dialogue, but it can nevertheless be seen that Habermas' theory provides good reasons for sanctioning teacher input into process over input into content.

Secondly, the teacher ought to exercise more power at the level of overall control of the educative environment, rather than at the level of determining each individual action. Creating an atmosphere in which students feel encouraged towards, and safe in, engaging in dialogue contributes more to the development of both reasonableness and autonomy than directing and immediately evaluating (in the IRF fashion) each person's input into the dialogue. Strategic action brings about behaviour in others that suits the actor's purpose. When this purpose is to induct the others into the autonomous use of reasonableness (an aim that is indexed to the needs of those others), then to maintain tight control over individual actions risks that the inductees, having never had to make the judgement about how and when to act, will never be able to take them over. Control over the educative environment, in order to be legitimate, needs to focus the (strategic) maintenance of dialogue in such a way that,

as far as the students are concerned, it approaches the standards of communicative action and/or critical discourse. This requires that the teacher play a dual role: participating in the dialogue whilst simultaneously monitoring it for opportunities to make the sort of intervention that model, encourage and develop good communicative moves and keeping the dialogue fruitful. Within such an environment, students can practice both reasonableness and autonomy.

These two observations are not unrelated: the maintenance of an educative environment depends critically on the presence of certain processes in the classroom. Nor are they clear cut. As I indicated in discussing the first point, process cannot be sharply distinguished from content. For example, process aims can be met by asking the right question of the right person at the right time, but what determines whether the question is the right one can depend as much on the content of the question as the process it encourages. With regard to the second point, the educative environment cannot be established independent of individual moves, for it is partially created and maintained in the encouragement or suppression of particular student actions, especially process moves.

4.141 Pedagogic Action

The analysis of the place of a teacher with respect to the three types of dialogical action advanced so far necessitates a closer look at how Habermas' categories apply to classroom talk, so as to address the pedagogic paradox outlined in §4.13. I have argued that the peculiar type of action that good teaching requires is a simultaneous and considered amalgamation of success-oriented action with the other two. But what is the nature of this amalgamation: is it a commingling of separate types of action, or are they fused into a single new mode? To accept this latter option is to posit a new category of Habermasian dialogical action: we could label it "Pedagogic Action." But ought we to posit pedagogic action as another type of dialogical action, to be placed alongside success-oriented action, communicative action and discourse, or would it be better to use pedagogic action as a label for the way in which Habermasian action types can be mixed? It may seem that there are reasons for preferring the latter. Habermas (1981) himself states that natural speech occasions can mix different types of dialogical action, while discussing the problem of:

distinguishing and identifying in natural situations action oriented to understanding and actions oriented to success. Here we must take into consideration that not only do illocutions appear in strategic-action contexts, but perlocutions appear in contexts of communicative action as well.... These strategic *elements* within a use of language oriented to reaching understanding can be distinguished from strategic *actions* through the fact that the entire sequence of a stretch of talk stands - on the part of all participants - under the presuppositions of communicative action. (331)

Relatedly, since Habermasian categories are closely linked to a theory of speech acts (see §4.121), a new category would seem to require the identification of a new type of speech act, if it were to fit within a Habermasian framework (Bob Young, pers. comm.).

Despite these considerations, however, I will argue that we ought to see pedagogic action as a further type of dialogical action. My argument turns on three points. Firstly, it is worth noting that Habermas is claiming that, although whole stretches of talk can be mixed, each stretch of talk nevertheless stands under the presuppositions of one type of dialogical action. He further claims that the “elements” (sentences? speech turns? - it is not clear what constitutes an element) of such a stretch can be individually allocated to dialogical action types. My analysis is different: in pedagogic action at least, single elements of talk can (and often do) *simultaneously* instantiate success-oriented action *and* (say) communicative action; they stand under, and can be analysed in terms of, the presuppositions of *both* types of action. We do not have a mix of strategic elements alternating with communicative elements; rather, single elements merge both. A question that asks a student to give reasons for their assertion simultaneously advances the inquiry at hand (communicative action) and assists in bringing about in the student a realization that reason-giving is a valuable move in discussion (success-oriented action). Both the element and the whole discussion are thus simultaneously success-oriented and communicative.

Secondly, although Habermas does align strategic action to perlocutionary speech acts and communicative action to illocution, he does not do so on a strict one-to-one basis:

I would like to suggest that we conceive perlocutions as a special class of strategic interactions in which illocutions are employed as means in teleological contexts of action.... A teleologically acting speaker has to achieve his illocutionary aim... without betraying his perlocutionary aim. This proviso lends to perlocutions the peculiarly asymmetrical character of concealed strategic actions. These are interactions in which at least one of the participants is acting strategically, while he deceives other participants regarding the fact that he is *not* satisfying the presuppositions under which illocutionary aims can normally be achieved.... I have called the type of interaction in which *all* participants harmonize their individual plans of action with one another and thus pursue their illocutionary aims *without reservation* ‘communicative action.’ (Habermas 1981, 294, italics in original)

All speech acts have locutionary, illocutionary and perlocutionary aspects. In pedagogic action, we do not have a situation of concealed strategic action - rather it is open strategic and communicative action merged. The teacher’s use of (say) Socratic irony does have perlocutionary aspects, but they are not essentially hidden. It does not matter to the achievement of the teacher’s aims that the perlocutionary aim is hidden, and often enough it will not be - the students may well understand that

the teacher's pretence not to know is a pretence. Not being illocutionary "without reservation," it cannot be merely communicative action, but not being perlocutionary in the concealed sense of wanting to achieve a result solely in the speaker's interests, nor is it mere strategic action. Thus we see the speech act markings of pedagogic action.

To make the final point, I need to sketch in some background. The phrase "pedagogic action" (German: *pädagogischen Handelns*) has been widely discussed in the continental literature, usually with explicit reference to Habermas and critical theory (see, for example, Masschelein 1991). The use of the word "pedagogic" here, rather than say "educative," points to the central intentional role of the teacher. A situation can be educative without the intention of any participant to make it so. Pedagogic action, on the other hand, essentially involves the intentional attempt by one participant (maybe a teacher, maybe a parent, maybe another) to educate the other(s). For a fuller analysis of the implications of the term "pedagogic" in the continental tradition, see van Manen (1991, 27-30, 40-48).

Kivelä (1998) points to some serious shortcomings in attempts (Masschelein 1991; Peukert 1993) to equate pedagogic action with communicative action (or communicative action plus critical discourse). This is the core of his argument: critical theory assumes that an ethical justification of education requires that dialogue in the classroom conform to the symmetrical character of communicative action, yet the Kohlbergian roots of Habermas' categories require that participants in communicative action have reached the post-conventional stage of reasoning. Empirical studies, however, show that such a stage is not reached until adolescence. Thus, it would seem that education is not possible until adolescence: a position which is obviously empirically incorrect.

I believe that, especially given the attacks on Kohlberg's work (and stage theory in general) from other sources (see §5.1 and §3.23 respectively), the Habermasian categories can be supported on grounds that do not entail post-conventional reasoning as a basis for communicative action. That task is too big for this thesis. Nevertheless, the more general point that Kivelä makes is sound: that any account of pedagogic action cannot assume that children are already competent in communicative action, that is, fully reasonable. It is important to be careful here with the word "competent." Clearly, in Habermas' ideal speech situation, all participants in discourse (and by extension, communicative action) are fully competent: i.e. perfect users of language. We need to remember, however, that this is a transcendental ideal, and hence counterfactual. Communicative competence in actual dialogical situations need not (indeed, cannot) reach these standards.

Communicative competence must be contextualized, like all claims of competence. For example, I am competent to drive the streets of Hobart, but not the Grand Prix track. Similarly, a kindergarten child may be communicatively competent within that community, but clearly not in the Oxford Union debates. Judgements of competence must be referred to standards, and standards depend on context.

There is a confusion at work here which I shall be addressing in more detail in §6.231. Firstly, “communicative competence” can be used to refer to the present ability level of the actor to engage in communication (cf my ability to drive on Hobart streets). Secondly, it can serve to mark the complete and normative communicative competence that would be counterfactually found in the ideal situation (cf my [non-existent] perfect ability to drive anywhere under any conditions). Finally, Habermas also uses competence to label the predispositions necessary to be able to develop the ability to communicate.⁴

Underlying this confusion, Biesta (1995) claims, are the presuppositions of the “philosophy of the subject,” which sees “the human subject as an autonomous, pre-social, trans-historical source of truth, rationality and identity” (274). Such an account makes it easy to lose sight of the fact that the details of the induction into communicative competence (in the first sense) will have marked effects on the rationality (I would prefer to say “reasonableness”) of the subject. It then becomes easy to talk, like Kant, of human subjects as “finite rational beings”: beings that have (amongst other things) achieved communicative rationality and competence (in the second sense). The next slide is to considering children as already having competence (in something like the third sense): (latent?) rationality and autonomy that just need maturity to emerge, with no place for a history of construction through social engagement. Thus the second and third senses merge, leaving no place for real people with real, variable reasonableness.

Depending on when we assume this miraculous transformation takes place, this picture of children risks either the conservative error of treating children as empty vessels that need filling with knowledge until they emerge into maturity, or the progressive error of treating them as small adults, who can gain the knowledge they need if only we let them exercise their autonomy in learning. Subject-based accounts of the pedagogical task of inducting the child into communicative competence are then prone to make three mistakes: they assume that the end point for the child to reach is ideal competence; that normal adults who take part in communicative action have already fully attained such ideal competence; and that

⁴ This explicitly follows Chomsky’s (1968) term “linguistic competence.”

any engagement in communicative action indicates (and possibly requires) this ideal competence.

The philosophy of the subject takes the actor as primary, and not the social situation in which the act (and hence the subjectivity of the actor) becomes possible. Biesta analyses this latter, alternative approach in Deweyan terms, though he points to the fact that Habermas and many postmodernist writers have also tackled this task of reconstituting subjectivity as intersubjectivity. Mind, according to Biesta's reading of Dewey, is not an original datum, but is acquired in the experience of communicative transactions (i.e. transactions imbued with meanings) with others. The young child is a proto-subject, in that it is biologically equipped to be able to take advantage of this social communicative experience and develop subjectivity, in a way that lower animals are not (§2.4). Nevertheless, it is only through the right sorts of experience that this development becomes possible, and the proto-subject develops into a becoming-subject, and finally the sort of situated subject, exhibiting a sufficient level of reasonableness and autonomy, that we characterize as an adult.

In abandoning the philosophy of the subject, we recognize that communication becomes, not just another task that the young subject (already rational, autonomous and individual) must learn to do, but a primitive interaction that already embroils the young child from the start, and is constitutive of the becoming-subject. Communication is a structure, a part of the lifeworld in which the child is always already entangled, while ideal communicative competence is a normative *telos*, a goal (counterfactual in its entirety) towards which the developing child can be aimed. It is the use of the phrase "communicatively competent" to label each of these alternatives which causes problems (§6.231). It makes sense to say that a normal neonate is communicatively competent in comparison to an autistic neonate *at the appropriate level*, but this is not to imply that the neonate is anything like a fully formed subject, autonomous and reasonable.

So, to claim as Kivelä does (correctly, in my opinion) that pedagogic action cannot be equivalent to communicative action between (normatively) communicatively competent actors is not to say that it cannot involve *any* element of communicative action (or critical discourse) at all. Children are certainly, by the time they get to school, competent to engage in communicative action at the appropriate level and given appropriate support (see, for example, Tizard and Hughes 1984 for empirical support for the pre-school claim). Meadows (1988, 30) comments:

It seems likely that... the difference between younger and older children will turn out to be that the former can do what the latter can; but only sometimes, only under favourable conditions, only with help, only without distractions, only up to a point,

without so much efficiency, without so much self-control, without so much awareness of the implications, without so much certainty.

As I have argued before on Vygotskian grounds (Sprod 1994b), this is precisely why, when given scaffolded support via social interaction to move into their Zone of Proximal Development, children are able to improve - to develop their level of reasonableness. Support offered through scaffolding is a crucial part of the process: without it, communication would remain at the same level. Social interaction in general leads to some development of reasonableness, for the sort of scaffolding we are talking about here is not unique to teachers. Adults, when interacting with children, commonly undertake such scaffolding intuitively, as demonstrated in Bruner's (1983) work with mothers and infants. Indeed, it happens informally in much social interaction of more with less competent actors. However, the fact that in general social dialogue, this support is often incidental to, and not a central focus of, such interaction means that the development of reasonableness in the absence of education is likely to be partial and piecemeal. These considerations lead us to the conclusion that to characterize pedagogic action as scaffolded communicative action and/or discourse is to make clear that it is a distinct type of dialogical action, for the scaffolding inherently involves open strategic action.

In summary, as I am characterizing it here, pedagogic action consists in a more experienced person interacting with a less experienced person with the specific aim of developing the latter's abilities. An important element of such pedagogic action, I maintain, is to develop the competencies to engage in communicative action or critical discourse, and this requires an inextricable and simultaneous merger of open strategic action with communicative action and/or discourse. Pedagogic action is clearly not confined to schools, or teachers with students, but is also engaged in by parents with their children, adepts with their apprentices and so on (van Manen 1991). It is marked by the use of power and some degree of coercion (such as requiring children to go to school), to achieve the often, but not essentially, hidden (from the children) aims of educating them. In a democracy, one of the central aims will be inducting the children into the wider speech community. Without this pedagogic use of power, the children will be unable to participate effectively in that community at a later stage.

The legitimacy of this success-oriented aspect to pedagogic action itself can, of course, be called into question by the students or by others. As Young (1992, 52) says: "The issue is whether the asymmetry present in a given situation is functionally justified and whether some safeguarding apparatus or process... is present." If such questioning, as a safeguarding apparatus, is not (at least sometimes) allowed to

become a part of the conversation, then the situation becomes inadmissibly distorted from pedagogical action. This is not to say that every challenge needs to be taken up at every stage - again, the teacher has power to wield in the interests of the children. Rather, such questions should enter the dialogue at appropriate times (as decided by the teacher), when consideration of them would serve to advance the children's induction. Thus the anarchic situation referred to by Blake (1995) is avoided. Moreover, the permitting and encouraging of critical questions is a crucial element in the endeavour to ensure that the teacher's pedagogic power does not lead to indoctrination (see §9.41). Such questioning can take place at many levels (within the classroom, at the school level, at the public policy level and so on); not all of which are open to the students. Still, to ensure that the strategic remains open, and thus the pedagogic action stays legitimate, there must be the possibility for students to question at a level appropriate to their developing competence and understanding.

Pedagogic action, under such an account, contains room for abuse. The power that the teacher has can be used to constrain the discussion and achieve aims (such as control and oppression) that are not educational. Yet we have to live with this possibility. The opposite situation, where the teacher abdicates power so as not to oppress the students, leads to abuse of a different sort - the denial of the child the chance to develop the capacities needed to be an empowered participant in the conversation of the community; in other words, the opportunity to develop reasonableness and autonomy. After a brief aside, I will turn again to autonomy.

4.15 Pedagogic Action in a Wider Context

In §4.1, I have suggested a way of reading Habermas on communicative action and of modifying his work to appropriately handle classroom dialogue. These conclusions could be extended in further work at two levels - the practical and the philosophical. In the first, my conception of pedagogic action shows promise for specific inquiries into classroom talk. It offers a way to think about the role of the teacher and the types of pedagogic techniques to be used, especially with an eye to improving the variety and quality of classroom dialogue. It also suggests that classroom research, particularly with a qualitative design, might be usefully guided through a theoretical framing in terms of different types of dialogical action.

At the second level, this discussion has merely hinted at several interesting general conclusions about Habermas' work. Firstly, it has cast doubt on the claim that dialogues, or even elements of dialogues, can be neatly categorized in the way Habermas claims, at least in educational contexts. By considering other settings, the question as to whether this applies more widely could be addressed. Relatedly, and

more importantly, it has questioned whether success-oriented action ought always to be characterized as involving merely the achievement of an individual's goals, for reaching such an end may be inextricably entwined with the achievement of socially desirable, legitimate goals for that individual's interlocutors. The interrelation of agents and their interests makes the neat separation of strategic action from communicative action suspect: in §7.12, we shall see that Habermas has been accused of drawing too-sharp distinctions in other areas. Finally, my work points to an important investigation of the role of the ideal and the counterfactual in the critique of ordinary everyday practice. While offering much promise, the use of such conceptualizations also raises much that is problematic. While I raise these points here, they are somewhat tangential to my main interest, so I shall drop them again and move on.

4.2 Autonomy Again

As we have seen, for Habermas what is rational is what the speech community (in the ideal speech situation) would decide is rational. It is a matter for negotiation: hence communicative rationality is inherently socially constructed, inherently part of being in an epistemic community. However, as in O'Neill's Kantian account of §3.221, we can again make a distinction between two levels: the first being the public level of historical rationality (the level on which Habermas concentrates), the other referring to the developing rationality of an individual. Of course, there are two constraints on the rationality of the individual: the finiteness of the reasoner and the variation from the ideal of their speech situation. In the historical context, the latter of these operates much more strongly than the former, because an epistemic community can, across all its resources, construct a model of rationality that exceeds the grasp of any individual in it (a point that is certainly true now, though it may have been much less so in earlier times).

While such a distinction can be drawn, the two types of rationality are related, for the individual builds rationality through interaction with many members of the epistemic community, and so has access to much of this wider version. Historical rationality, however, is in many ways a fiction: it supervenes on the actual individual rationalities of all the individual members of the community, and all access to it is through such individual members. I shall return to this point in discussing the problems for Habermas' account.

So we now have a picture of the child as an apprentice building reasonableness, firstly by just listening to others, then by beginning to take part, as more competent users scaffold the development, and finally by internalising the competencies

sufficiently to be able to “play the game” on their own. The child is becoming more autonomous:

Habermas’ communicative theory of epistemic justification provides the content of a concept of autonomy as the ability to participate in argumentation (communicative competence). (Braaten 1995, 141)

An analogy, drawing on the phrase “play the game” in the previous paragraph, may make this process clearer. A young girl, of say two years, is pretty well incapable of playing basketball. Certainly she has a basis on which to build: she can walk and catch a soft ball if it is not thrown too hard. Also, she can watch others play and try to imitate them. As she grows, and as others include her in their games (but carefully, modifying the speed, the physical demands and probably the equipment used in the game), she is, with help, able to do more and more. By say the age of 12, she is able to play a game that is recognisably real basketball - but she is still not very good at it. Her variations from the standard, accepted ways of playing basketball are much more likely to be due to lack of skill and understanding than to inspired innovation, though there are occasional flashes of the latter. In a game with adults (unless she is remarkably exceptional), she would not be able to match their speed, strength nor, importantly, their ability to read the game and make important strategic decisions well.

In training, she continues to listen closely to her coach, to drill the fundamentals of the game and to watch older and more experienced players. This drilling is essential: she must learn to make these skills (in Polanyi’s terms, §3.1242) a part of her subsidiary awareness of the game, rather than part of her focal awareness. It is in this learning to be able to carry out the moves in *just the same way* as every other highly skilled player that she becomes eventually able to be an outstanding player herself. Innovation, foresight, originality and independence in play *depend on* the internalization of the common skills of the game, rather than springing out of an independence from them. Similarly, to be a great player requires the ability to blend those skills seamlessly with the other players in the team: the “brilliant” pass that the receiver can’t handle is not brilliant at all. In similar style, the autonomous reasoner emerges from immersion in the basics of reasoning, held by all reasoners alike.

Nevertheless, there are problems with the Habermasian account. While Habermas does deal with the development of communicative competence, for the most part in his work the reasoners that enter into dialogue are adults, who have already constructed reason. The account he does give of development rests heavily on the work of Piaget and especially Kohlberg (Habermas 1990, 33-41, 119-189; Kivelä 1998, see also §4.141), but this provides a fairly thin account. Piaget neglects the

social and interactive dimension of development, drawing overmuch on the interaction of the reasoner with the physical environment (§3.231). Kohlberg develops an account of moral development that concentrates narrowly on justice reasoning based on universal principles, over-emphasizing argumentative rationality and neglecting the roles of the affective, the habitual formation of virtues and paradigm scenarios (§5.1: see also Gilligan 1982; Matthews 1994).

The works of Piaget, Kohlberg and Habermas all have explicitly acknowledged Kantian roots. In each, formal principles take precedence over content. As we have seen (§3.21), for Kant, the Moral Law is an abstract, formal principle, legislated solely by a formalistic Reason, that determines morality. Piaget casts all development as transformation of the formal logical principles underlying thought, which are generally applicable across content domains (Donaldson 1978, 133-140). Kohlberg explicitly takes Kant's moral theory as the highest stage of moral development (Kohlberg, Levine et al. 1983). Habermas (as we shall see in §6.1) claims that two formal principles, empty of normative content, define his discourse ethics. These principles (as do Piaget's and Kohlberg's) draw explicitly on Kant's notion of universalizability.

In §6.24, I shall argue (specifically against Habermas, but drawing on the Hegelian critique of Kant), that such formal principles are insufficient to establish substantive norms, being devoid of content. My account of reasonableness (§3.1) is much thicker than such content-free descriptions, because connection to real people in real communities requires the inclusion of contextual and situated factors such as social locatedness, emotionality and embodiedness.

Similarly, Habermas conceptualizes reason as a fairly narrow rationality, as I shall further explore in §6.23. This concentration on rationality contributes to his acceptance of the possibility of full consensus in the ideal speech situation, an acceptance that would be much more problematic if a rich concept of reasonableness, as I have previously outlined, was to be accepted. Finally, the reasoners in the Habermasian ideal speech situation are often characterised as rational beings, with differences between them downplayed. As Benhabib (1992, 58-70) has emphasized, the dialogical partners in Habermasian communication are conceptualized as the Generalized Other, rather than as specific Concrete Others who differ greatly in many respects. Habermas has attempted to escape from the Kantian insistence, seen also in Rawls' work, that a monological subject can make moral decisions binding on all (I discuss this at length in §6.1). However, as I shall also discuss (§6.232), his insistence on the inevitability of a consensus in the ideal speech situation means that differences must be set aside.

Thus, while we have the core of a reconceptualized autonomy here, Habermas' account needs to be extended: it needs to take more fully into account the formation of individual reasoners in diverse situations that have real differences between them. This point has been made by a number of feminist and communitarian commentators, who appreciate the power of Habermas' recognition of community to deal with real, situated and embodied persons.

4.21 Dependence throughout Development

It is tempting to simplify any philosophical task involving persons by taking oneself as the paradigm person. Perhaps the most famous example of this is Descartes' (1984, 17) "Cogito, ergo sum," embedded as it is in the introspective narrative of the *Meditations*. If the endpoint of development is taken as something universal, that every person eventually reaches, then this tactic will be fruitful. But it is obvious that as adults we all have idiosyncratic features which differentiate us from one another, as well as sharing many features in common. There are two philosophically important points here: first, whether these differences are sufficiently unimportant, and the similarities sufficiently robust and wide ranging, that a simplification of the complexity of all persons to a single idealized exemplar will not lose anything of philosophical interest; and second, whether the pathway by which such adulthood is reached makes any significant difference to the final product.

Feminist philosophy has led the commentary on the first point. Many writers in this tradition have pointed out that the position of advantaged white males in the world is importantly different from that of others, especially women (Lloyd 1984). Gilligan's work (1982) on the ethics of justice and the ethics of care, tied more closely to males and females respectively, is a landmark of this type. More recent feminist scholarship (Spelman 1988) has called into question whether we can even use categories such as "women" without erasing important differences within the category (for a critique of Habermas from this perspective, see Warnke 1995). I shall expand on this aspect later.

The second point has not received as much attention, to my mind. Some feminist and communitarian writers have made some contributions, as I will discuss below. Nevertheless, a more developed philosophy of childhood will be able to highlight important philosophical implications from the development of children for mainstream philosophical debates in the same way that feminist philosophy has had major impacts in many traditional fields of philosophy. In this section, I will sketch out some of the grounds for that claim.

Firstly, however, let's see what we can glean from the writers alluded to above. One of the chief contributions of feminism is that:

Feminist epistemology tends to argue that the subject comes to the specialized activity of justifying knowledge claims as an already socially embedded being. (Braaten 1995, 150)

A seminal (or should that read “ovarious”?) text is Baier's (1985) essay *Cartesian Persons*, in which she points out that in developmental terms the first person is not chronologically first. The sort of dependence on others for the formation of the self, reasonableness and the capacity to operate in the world that I have already discussed (§3.23) leads Baier to say that we are all at base “second person”:

Self-consciousness depends upon exercise of the cultural skills, in particular linguistic ones, acquired during our drawn-out dependency on other persons. A person, perhaps, is best seen as one who was long enough dependent on other persons to acquire the essential arts of personhood. Persons essentially are *second* persons, who grow up with other persons. This way of looking at persons makes it essential to them that they have successive periods of infancy, childhood, and youth, during which they develop as persons. (84)

This way of looking at persons makes it essential to pay serious attention to what goes on during this period of “long drawn-out dependency.” In my case, this means attention to the development of reasonableness and of the ability to function autonomously, two of the “essential arts of personhood.”

Code, with an epistemic focus, has claimed that one important paradigmatic notion of “knowing” lies in the knowing of another person, in friendship, as opposed to the solitary knowing of an inanimate object sitting in front of us. She criticizes the assumption that knowledge as a product is independent of the process, and hence independent of persons, claiming that, without knowledge of persons, knowledge of objects becomes deeply problematic. If we do pay attention to Baier's “successive periods,” especially infancy, we find that there is empirical evidence to support Code's assertion.

Many philosophers have noted that objects are relatively simple when compared to persons. From this observation they draw the conclusion that epistemological investigation should start with examples such as the perception of red boxes. However, developmental studies have shown that young children's knowledge of persons develops earlier than, and far outweighs, their knowledge of even simple objects. Infants attend preferentially to faces, and especially the eyes, over other similar shapes from birth (Karmiloff-Smith 1992, 19; Baron-Cohen 1995, 40). A child about to enter school will use misconceived mechanical ideas to wrongly predict the motion of even simple objects in simple situations (Yates, Bessman et al. 1988; Millar and Kragh 1994), but will also be able to make accurate predictions of

the actions of people (e.g., by ascribing false beliefs to them, Baron-Cohen 1995, 69-72). Infants and children do seem to be able to grasp knowledge of persons earlier and in more detail than knowledge of objects. The basis for much of this is built-in to our cognitive architecture. Cosmides and Tooby (1994, 53-54) say:

the adaptive problems posed by social life loom large. Most of these are characterized by strict evolvability constraints, which could only be satisfied by cognitive programs that are specialized for reasoning about the social world. This suggests that our evolved mental architecture contains a large and intricate 'faculty' of social cognition.

The development of these social capabilities, however, depends on interaction in the sociocultural environment (Karmiloff-Smith 1992, 122). The gaining of intersubjective knowledge and capabilities through social interaction underpins, in large part, the ability to gain knowledge direct from the world in interaction with it. So to take the already adult solitary rational knower as the basis for a philosophical investigation into the nature of autonomy is to start well away from the roots. An adequate concept of the reasoner upon which to base autonomy must take account of the formation of reasoners rather than take reason as given.

As I have been arguing, the formation of reasoners is an inherently social process, involving development that starts with a largely incapable infant, full of potential which can only be realized through interaction and dialogue with others. Benhabib (1992, 5) captures this well:

The subjects of reason are finite, embodied and fragile creatures, and not disembodied cogitos or abstract unities of transcendental apperception to which may belong one or more bodies... the subject of reason is a human infant whose body can only be kept alive, whose needs can only be satisfied, and whose self can only develop within the human community into which it is born. The human infant becomes a 'self,' a being capable of speech and action, only by learning to interact in a human community. The self becomes an individual in that it becomes a 'social' being capable of language, interaction and cognition. The identity of the self is constituted by a narrative unity, which integrates what 'I' can do, have done and will accomplish with what you expect of 'me,' interpret my acts and intentions to mean, wish for me in the future etc.

Of course, it is not only in infancy that we define ourselves and gain our powers through interaction. Taylor (1989, 36) points out that this dependence lasts a lifetime. Although we become progressively more able to act on our own, this is only achieved, and is always supported, through that interdependence:

Even as the most independent adult, there are moments when I cannot clarify what I feel until I clarify it with certain special partner(s), who know me, or have wisdom, or with whom I have an affinity. This incapacity is a mere shadow of the one the child experiences... In this sense, one cannot be a self on one's own. I am a self only in relation to certain interlocutors: in one way in relation to those conversation partners who were essential to my achieving self-definition; in another in relation to those who are now crucial to my continuing grasp of languages of self-understanding - and, of course, these classes may overlap. A self only exists within what I call 'webs of interlocation'. (36)

This developmental picture of autonomy portrays persons who become increasingly able to function more independently whilst never losing the engagement with others that enables such independence. And while engagement can take many different forms, Code (1991, 121-2) agrees with Taylor's emphasis on interlocution, giving the central place in the construction of thinking and knowing selves to dialogue within a community.

Epistemological positions developed around a 'second person' conception of subjectivity represent the production of knowledge as a communal, often cooperative though sometimes competitive, activity. Either way, knowledge claims are forms of address, speech acts, moments in dialogue that assume and indeed rely on the participation of (an)other subject(s), a conversational group... dialogue is primary, so that even thinking, that seemingly solitary activity, is constructed on a conversational model.

Drawing on her discussion of the importance of testimony in what we can know, a point I have discussed in §3.2413 in relation to thinking for oneself, Code approvingly quotes Hardwick as

claiming that the Peircian conception of the *community* of inquirers as the primary knower, in which individual knowledge is derivative, is the most plausible one, in view of the role of circumspect reliance on testimony as a *condition* of responsible knowing. (Code 1991, 132)

This reliance on the testimony of others is not solely, or even primarily, a reliance on others for our stock of facts. As Code recognized in the passage previously quoted, thinking itself depends on the testimony of others - on their modelling and scaffolding of the very flesh and bones of reasonableness itself. Pettit (1993) argues that it is not even merely the higher tools of thinking that are gained through social interaction, so that we start with innate rudimentary conceptual thinking and refine it socially. Rather, it is the possibility of turning preconceptual, merely representational mentation (Pettit's non-thinking intentionality) into conceptual thinking that depends on immersion in a linguistic community. Further philosophical support for closely related conclusions can be drawn for the works of Wittgenstein (1953) in the famous private language argument, Sellars (1963) and Mead's (1982) argument that reflective intelligence is possible only because of the social component of mind. We saw the empirical and psychological support for these philosophical conclusions in §3.23, where the Vygotskian project was described.

So far my argument has been that infants, dependent on others for so much, have to rely on others for the development of their reasonableness. This reliance is not passive but interactive. As children grow, and as they participate more and more in dialogue that jointly constructs reasoned arguments and discussions, they become

equipped with the attributes of reasonableness, together with increasing knowledge of the world. All this has implications for the sort of autonomy they are building.

4.22 Communicative Autonomy

The term “autonomy,” as we have seen, has many connotations. Its Kantian lineage paints the notion of autonomy as austere, separated and individualistic. Yet I have argued that a concept of autonomy must have a place in our consideration of persons as moral and ethical agents. This reconceptualized notion cannot be based on a strong sense of independence of persons one from another. It has to recognize and celebrate the intersubjective interdependence which makes us capable of autonomy. Hence, I propose to label it *communicative autonomy*.

I will not make the claim that communicative autonomy covers all the areas previously covered by the Kantian version of autonomy. In particular, communicative autonomy is not taken to cover actions, as distinct from the judgements that lead to actions. Nevertheless, such act autonomy is, I claim, parasitic upon communicative autonomy. Actions (I am making a distinction here between speech acts and physical acts and referring largely to the latter), if they are not to be random, issue from judgements. Judgements, in turn, issue from dialogue: either the overt dialogue that we have when we discuss our options with others, or the internalized dialogue that we call thinking.

Communicative autonomy is the ability of the self to engage fully in the ongoing conversational narrative of humanity. Of course, as life-long learners, we never reach that idealized ability to participate perfectly. Our rationality is always limited, our imaginations are bounded, our emotional judgements can interfere with as well as support our thinking, we are contextually situated in ways that restrict our thinking and we are restrained by our embodiment as individual persons.⁵ Nevertheless, as our abilities to expand our horizons grow, through engagement with the differing horizons of others, so we become more autonomous.

Through the internalization of socially mediated practices, including the practice of critique, we become more able to engage reflexively with others and with issues. Communicatively autonomous persons in this characterization are persons who have appropriated the standard, intersubjectively shared rules, prescriptions, techniques and capacities of reasonableness well enough to be able to wield them according to their own beliefs, desires, aims and goals. They can modify their beliefs, desires,

⁵ It should be noted that the constraints placed on our thinking can both hinder and help our thinking. Constraints do the latter by restricting the incoming data that will be processed, and by cutting down the possibility space in which we have to make decisions. Given time and processing limitations, this enables us to make reasonable decisions in real time (Karmiloff-Smith 1992, 11).

aims and goals in the light of reasonable grounds for doing so. They can show flair and originality, they can adapt them to different purposes at need. They can apply their reasonableness uniquely, when a unique application is needed, but they can also use it in standard ways when that is adequate to the task.

Such autonomy does not mark the holder off from others. As the development of reasonableness is inherently intersubjective, so is the formation of needs and aims. Being immersed in communicative situations, the increasingly autonomous agent relies on Taylor's "webs of interlocution" to build and to exercise autonomy.

Autonomous agents are not identical, for the limits of communicative autonomy are different for each, as is the specific personal version of reasonableness upon which it rests. Communicatively autonomous persons thus differ in respect of both their autonomy and their reasonableness. Indeed, we cannot even put persons on a continuum and say that each has developed autonomy to a particular degree. Because autonomy rests in large part upon reasonableness, and because reasonableness is a multi-aspectual capacity, then comparisons of one person with another in regard to their development of autonomy are not straightforward. Indeed, the many other factors that contribute to the notion of communicative autonomy (such as listening skills, verbal fluency, personality traits like shyness and so on) complicate the picture even more.

A multi-aspect theory of reasonableness implies that different people may develop greater abilities in one or more of the aspects of reasonableness than others. The claim here is similar to the one made by Gardner (1985) with reference to his multiple intelligences theory. Gardner denies that it makes any sense to claim that people can be put on a continuum of intelligence, such as that offered by IQ tests. If there are seven distinct types of intelligence, each with some interactional effects with the others, then efforts to reduce them all to a single figure are meaningless. Similarly, we cannot talk as if there is any meaningful way to rank someone with a vivid, inventive and focused imagination but poor logical skills against a person with a mind like a steel trap, but a lack of imagination. The contextual aspect in thinking also means that it is quite possible that the one and the same person will be able to join in to some conversations with a high degree of autonomy, yet display considerable heteronomy if the situation is changed. Gilligan's (1982) research shows that, on average, the females in her samples tend to be more comfortable using reasoning with a greater emphasis on the committed aspect of reasoning, while the males tended to focus more on the critical. Her case studies also showed that not all males and females fitted this pattern, and that changes in the context of the

discussion would often lead to a change in emphasis from the committed to the critical or vice versa in the same subject.

4.3 From Foundations to Moral Theory

In Part I of this thesis, I have philosophically explored two major aims of the educational endeavour: reasonableness and autonomy. I have suggested that the key to reconceptualizing rationality as reasonableness, and autonomy as communicative autonomy, has been to take development seriously. When this is done, we see that reasonableness cannot be a singular and identical capacity in all humans, while autonomy cannot deny interdependence.

Through most of this survey, I have not concentrated specifically on the implications of these two capabilities for moral education. Before I can do that, I will need to outline a moral theory that establishes an aim for moral education. This will be the task of Part II. Once again, we shall see that a lack of attention to the developmental trajectory of moral development has resulted in some philosophical weaknesses in otherwise sound moral theories.

Part II: Moral Theory and Moral Development

Chapter Five: A Target for Moral Education

It is pointless to talk about a moral education program if it is not made clear what counts as morality. Further, an educational program needs to have a clear idea not only of its endpoint, but also of the developmental path that leads towards that endpoint. In this Chapter, I shall sketch a meta-ethical account that meets these two requirements. To construct and defend an entire meta-ethical system would be far beyond the scope of this thesis, so I shall take the approach of contrasting two accounts of moral development, identifying the meta-ethical theories that underlie them, and then developing an account of a plausible goal for moral education that draws on elements from these two theories.

5.1 Moral Development and Meta-ethical Theory

The most influential theory of moral development this century has been that of Lawrence Kohlberg (see, for example, Kohlberg, Levine et al. 1983; Power, Higgins et al. 1989, also §4.2). Kohlberg sees moral development as development along a single dimension - that of moral judgement. His theory has two roots: the genetic epistemology of Jean Piaget (1977) and Kant's meta-ethical theory, via Rawls (1973). Piaget explicitly states in his Foreword that "it is the moral judgement that we propose to investigate, no [sic] moral behaviour or sentiments," thus leaving open the question as to whether his studies encompass all of moral development or only a part of it. Kohlberg, on the other hand, rules out any other component of moral development. His theory asserts that justice is a universal moral principle, that it can be taught, and that all else flows from it (Power, Higgins et al. 1989, in the Chapter From Moral Discussion to Democratic Governance, written with Joseph Reimer).

However, the universality and completeness of Kohlberg's theory has been called into question by many. Perhaps the most celebrated attack came from within the Kohlberg camp. Gilligan (1982), in attempting to replicate Kohlberg's results (originally obtained using a cohort of males only) with young females, asserted that the Kohlbergian stages left out an important second dimension of moral development - that of care. Other attacks have drawn on a number of other points. Kohlberg himself (1983) has summarized some common threads: critics assert that his account applies a normative theory as universal; misrepresents those with different perspectives; and is incomplete in failing to take into account other factors

than reasoning about justice (e.g. imagination, affect, sense of responsibility). Let's take a closer look at one of these critics.

Matthews (1994, Chapter 5), following an extended critique of Kohlberg's account, identifies five dimensions of moral development, acknowledging that these may not exhaust the field. Matthews' elements are:

- M1. a situated, experiential stock (or knowledge) of moral paradigms, which is gradually enriched with further experience, that forms part of the base for moral intuitions, against which the other dimensions may be measured;
- M2. an increasing ability to be able to offer defining characteristics for moral terms that take account of their complexity;
- M3. an increasing ability to judge whether a range of cases fall under a particular moral term, especially borderline cases;
- M4. a growing sophistication in the adjudication of apparently conflicting claims, when moral intuitions collide;
- M5. a heightening of the moral imagination, based in part on increasing understanding of the world and how it works.

Matthews (*ibid.*, 60) points out that Kohlberg concentrates solely on the fourth of these dimensions. Since in his scheme, pre-conventional (Stage 1 and 2) moral reasoning is purely prudential, and conventional (Stage 3 and 4) moral reasoning is merely conformist, Kohlberg's account seems to deny that anyone can really act for moral reasons until they reach Stage 5 or 6. Yet most of the population, according to a wide range of Kohlbergian empirical studies, never reach post-conventional reasoning, and hence, must be said not to have a grasp of what morality really is. This seems to be counter to our experience, in which even quite young children do at times act for reasons that we would want to call moral. A multi-dimensional account of morality allows us to do so, on the basis of one or more of the other dimensions of morality.

Johnson (1993) also advances a multi-dimensional account of moral development. He appeals to the concept of the moral imagination, which itself has multiple aspects. A doctrine of Universal Reason, like Kant's (§3.21 and §3.22), "ignores the crucial role in our moral understanding that is played by our bodily experience, our emotions, our imagination and our interpersonal and cultural relations" (*ibid.*,

245). This rich characterization of moral imagination clearly goes beyond merely including a creative aspect into critical thinking (as the term “imagination” might lead us to think). Johnson’s moral imagination explicitly includes embodied, committed and contextual elements, as well as the creative and critical (to use my terminology).

Johnson lists five challenges to what he characterizes as Moral Law theories (such as Kant’s and thus Kohlberg’s), which arise from recent empirical findings in cognitive science:

- J1. The replacement of a “necessary and sufficient conditions” characterization of concepts and categories with a theory of prototypes, so that moral concepts now must be seen as having core meanings surrounded by fuzzy grey areas;
- J2. Frame semantics, whereby terms get their meanings only within a larger context, implying that contextual settings are of vital importance to moral matters, and that different people might frame what appears to be the same situation in different ways;
- J3. The centrality of metaphor to understanding, so that moral (and other abstract) situations are understood in terms of mappings onto more concrete, maybe embodied, domains. This also introduces plurality into moral understandings, as well as providing a metaphorical extension mechanism for extending our moral construal of situations;
- J4. Basic-level experiences (often embodied, and hence universal) have a priority in the way we frame moral situations, so that there is a way to introduce some degree of universality into the moral domain without relying on universal moral laws;
- J5. The centrality of narrative as our form of understanding and experiencing moral situations, which provides another mechanism for extending moral understanding.

It can be seen that, despite differences of emphasis, Johnson’s points underwrite Matthews’ account in a number of ways. I shall not try to tease them all out, but we can see that the narrative form of understanding (J5), combined with frame semantics (J2), underlie Matthews’ situated and experiential stock of moral paradigms (M1), while the indeterminacy provided by the prototypical structure (J1) and metaphorical nature of concepts (J3) makes sense of Matthews’ assertions that

children need to gradually develop the ability to define moral terms (M2) and judge when cases fall under those terms (M3).

Multi-dimensional accounts of moral development are often (as in the case of Matthews') based on a virtue ethics such as that presented by Aristotle. I shall also base my account of moral development on a broadly Aristotelian foundation, and the next few subsections will explore Aristotle's views and their relevance to this thesis. Firstly, however, I shall look a little more closely at the advantages of multi-dimensional accounts of moral development over a Kohlbergian account.

The final stage of moral development (Stage 6), according to Kohlberg, is the principled stage. When people reach this stage (and it must be borne in mind, as previously mentioned, that very few people have been found to reach it in empirical studies), they realize that all moral judgements must be based on the universal principle of justice. Given that there is a single principle of judgement, all moral questions are, at least in theory, decidable, although practical difficulties may make it hard to find the universally acceptable answer. Such a moral theory would, if true, meet Mill's (1965, 278) criterion that:

there ought either to be some one fundamental principle or law, at the root of all morality, or if there be several, there should be a determinate order of precedence among them; and the one principle, or the rule for deciding between the various principles when they conflict, ought to be self-evident.

There are problems with asserting that all moral problems are in principle decidable, as I shall discuss in more detail in §6.2322, where I consider Habermas' discourse ethics. A multi-dimensional account, such as Matthews suggests, allows the consideration a number of ethical dimensions in moral judgement. But, in the absence of Mill's "determinate order of precedence among them," then there is no algorithm for deciding between these dimensions, if they are in conflict. This might seem to be a disadvantage in comparison to a Millian theory, as it would mean that there is no sure way to decide whether an action is morally acceptable or a person is good. However, I argue that this is, in fact, an advantage for multi-dimensional accounts, as it captures what is surely one of the most striking features of moral discourse - the fact that agreement on these matters is difficult to reach.

In her book *Postures of the Mind*, Baier (1985, particularly in the essays "Theory and Reflective Practices" and "Doing without Moral Theory?") argues the advantages of multi-dimensional accounts over Millian theories (without using this particular terminology). She claims that:

A significant fact about moral conscience is that its deliverances need not come in verbal form, that it is often a difficult task to articulate what it is that we are certain

is wrong in an action, let alone what universal rule it breaks. In moral philosophy courses we insist that students make their moral intuitions articulate, that they represent them and “defend” them by subsuming them under some universal rule that coheres in some system, and we make them feel that they must have been muddled if their moral intuitions are inarticulate or resist tidy codification. But it may be we the intellectualizers who are muddled... (213)

Baier reminds us that any account of what it is to be moral must take into account the phenomenology of moral thinking; it needs to be sensitive to our everyday experience of moral action. While it could be true that such inarticulate moral experience merely betrays a lack of sufficiently explicit consideration of the moral domain, it could also be the case that moral intuitions are not capturable in explicit formulations. A multi-dimensional account is consistent with this moral phenomenology, as it allows for various diverse sources of moral input. Due to the nature of their source, some of these may not be easily articulable. Similarly, if they spring from different sources, there may be real problems in merging them in a way that is linguistically reportable. Any meta-ethical theory that fits Mill’s criteria, on the other hand, seems to imply a high degree of linguistic explicitness, both in the statement of the principle(s), and in the deduction of implications.

A Kantian meta-ethical system is based solely on reason. Since in such accounts reason is linguistic and unitary, it is not surprising that the Kantian theory underlying Kohlberg’s work fits Mill’s criteria. A multi-dimensional theory, on the other hand, need not eschew reason (provided it is construed more widely as reasonableness) as one of the sources for morality, but it does allow that other sources can contribute. In what follows, I shall assert at least three sources of morality: the genetic inheritance of humanity (Chapter Two), the social situations in which humans are raised (§3.23), and the influence of reason (§3.1). Aristotle, as we shall see, also makes provision for each of these three sources. It is time to turn to Aristotle in order to clarify what a multi-dimensional account of meta-ethics might be like.

5.2 Aristotle’s Virtue Ethics as a Multi-dimensional Meta-ethical Theory

In this section, I will consider Aristotle’s virtue ethics as an example of a multi-dimensional meta-ethical theory. In doing so, I will illustrate some of the key features of multi-dimensional accounts and explore the extent to which Aristotle’s version is adequate. We will find that while there is much in Aristotle that is useful, he says little about one of the features most important to the thrust of this thesis: the place of dialogue in moral education.

According to Aristotle, what is most important to morality is neither following an obligation nor bringing about the best consequences, but being a good or virtuous person. Being a good or a virtuous person is an important aspect of achieving the

inherent aim of human life - *eudaimonia* (I.7¹, translated by Ross as happiness, but often rendered as “well-being” or “human flourishing”). Aristotle writes (I.13) that there are two groups of virtues: intellectual (e.g. theoretical wisdom, understanding, practical wisdom) and moral (e.g. temperance, liberality). Thus there are a variety of virtues, and the question arises as to how the virtues are coordinated in specific moral situations. Before I consider this, however, there is a need for a closer look at what the virtues are and how they are acquired. Here, then, is Aristotle’s definition:

Virtue, then, is a state of character concerned with choice, lying in a mean, i.e. the mean relative to us, this being determined by a rational principle, and by that principle by which the man of practical wisdom would determine it. (II.6, 1006b36)

In this definition, there are several key points to note. Firstly, virtue is about character, a relatively stable propensity to choose to act in certain ways. Secondly, virtue chooses a middle way (Aristotle’s Golden Mean), and since the determination of that middle way is relative to us, the mean is not something that can be decided by reference to principles only. Elsewhere, Aristotle emphasises the situatedness and lack of deductive precision of ethical decisions as follows:

The whole accounts of matters of conduct must be given in outline and not precisely... accounts we demand must be in accordance with the subject matter; matters concerned with conduct and questions of what is good for us have no fixity... the agents themselves must in each case consider what is appropriate to the occasion. (II.2, 1104a1)

Thirdly, Aristotle’s account of virtue involves rational principles as determined by practical wisdom - and the latter is clearly embodied in a person. We need an account of rationality that will underlie the virtues. I believe that the account of reasonableness developed in §3.1 is admirably suited to this role, because its multi-aspectual character suits a multi-dimensional moral theory.

As to how the virtues are acquired, there is a certain tension in Aristotle’s account. I shall explore this tension by means of a commentary on several passages:

Neither by nature, then, nor contrary to nature do the virtues arise in us; rather we are adapted by nature to receive them, and are made perfect by habit. (II.1, 1103a25)

From this, it is clear that the virtues are not fixed in us at birth - they are not merely natural reactions. Nevertheless, at birth, we do have an ability to be able to develop them (§2.4). Humans are in some way adapted so as to be likely to become virtuous, although in the absence of the formation of habits, they will not develop. Note that in this passage, Aristotle claims that the virtues are *made perfect* by habit. This

¹ My exposition here draws on the Ross translation of *The Nicomachean Ethics* (Aristotle 1980). While I shall often cite the book and section number in the form used by Ross (e.g. Book Two, section 6 as II.6), reference to more specific citations and quotations will be in the standard form, referring to the Bekker text (Prussian Academy, Berlin, 1831) by page, column and first line of the quotation.

would seem to imply that it is habit alone that develops the virtues. Here he expands on what it is to form a habit:

Thus, in one word, states of character arise out of like activities. This is why the activities that we exhibit must be of a certain kind; it is because the states of character correspond to the differences between these. It makes no small difference, then, whether we form habits of one kind or another from our very youth; it makes a very great difference, or rather *all* the difference. (II.1, 1103b19)

The formation of habits, then, takes place through repeated activities that are in conformity with the virtues. We must again and again act as the virtue would require if we are to form the habit. He claims that this is an important process: it make *all* the difference as to whether we can become virtuous or not. So far, the account that seems to be building is one that claims virtue depends solely on habit. Two things are not yet clear: where the decision to form one habit rather than another comes from; and the place of reason in the acquisition of habits. I shall leave consideration of the second to one side for a while.

Given that Aristotle is referring to habit forming “from our very youth,” it might seem that the sole source of these habits must be their imposition on us from the outside, by our parents, perhaps (though Aristotle also argues that the law should be involved). I shall refer to this view as mechanical habituation. What it requires - and this has been a common interpretation of Aristotle, as I shall document in §5.21 - is that parents, or society, tell children how to act in certain morally charged situations, and children just do as they are told, until eventually the moral action is “second nature.” Only when a person has been trained to act in a virtuous manner unflinchingly is it possible to build true virtue through the study of moral philosophy.

If this is the whole story, then the sorts of habits we form will be beyond our control. Yet, there are other passages in the *Nicomachean Ethics* that cast doubt on this interpretation:

All who are not maimed as regards their potentiality for virtue may win it by a certain kind of study and care. (I.9, 1099b20)

But perhaps a man is the kind of man not to take care. Still they are themselves by their slack lives responsible for becoming men of that kind, and men are themselves responsible for being unjust or self-indulgent, in that they cheat or spend their time in drinking-bouts and the like; for it is activities exercised on particular objects that make the corresponding character.... Now not to know that it is from the exercise of activities on particular objects that states of character are produced is the mark of a thoroughly senseless person. Again, it is irrational to suppose that a man who acts unjustly does not wish to become an unjust man... But if *without* being ignorant a man does the things that will make him unjust, he will be unjust voluntarily. Yet it does not follow that if he wishes he will cease to be unjust. (III.5, 1114a3)

These passages make it clear that the development of habits is under a certain degree of voluntary control. Virtue can be attained by “a certain kind of study and care.”

Just what kind, is identified in the second passage. We make the choice to engage in certain activities. If these activities are the sort that are “unjust or self-indulgent,” then we reinforce our unvirtuous habits. Nevertheless, we can, by taking care, provided that we are not ignorant, recognize this and choose to act differently. Aristotle is not naive enough to suggest that a single choice of this nature is enough to reverse an unvirtuous habit: ceasing to be unjust requires the training of a new virtuous habit, and this implies that we must continue to make the explicit choice to act virtuously until it becomes a habitual reaction. How are these choices to be made?

Since moral virtue is a state of character concerned with choice, and choice is deliberate desire, therefore both the reasoning must be true and the desire right, if the choice is to be good, and the latter must pursue just what the former asserts. (VI.2, 1139a22)

Deliberation, using practical reasoning, is the key to such choice. When we are not acting habitually, we deliberate about what the right choice would be. It seems that this is where Aristotle sees the need for care. If we care about what sorts of persons we are (see §3.1235 for an account of “caring about”), then we monitor our habitual reactions, as they are formed at the moment. If we suspect that these habits are unvirtuous, then we need to invoke practical wisdom so as to deliberate about whether to change our habitual reactions. Clearly, the committed aspect of reasonableness plays an important role in practical reason here.

Now, these reactions have been formed by our actions in the past, and many of these actions could well have been those that others (our parents, to a large degree) have required of us. We are situated in the world in a way that depends on our past. But we do not escape responsibility for our reasoning and our desires on this account. Our moral character is never a finished product: we are always “becoming men of that kind.” The kind we are becoming is under our control - not an easy control, but a control nevertheless, and one which depends critically on deliberation and reflection using practical wisdom.

But this is not to exhaust the role of practical wisdom. It also plays an important role in regulating the application of the virtues, which are general dispositions to act, in specific ethical situations.

But to feel [the virtues] at the right time, with reference to the right objects, towards the right people, with the right motive, and in the right way, is what is both intermediate and best, and this is the character of virtue. (II.6, 1106b24)

The work of man is achieved only in accordance with practical wisdom as well as with moral virtue; for virtue makes us aim at the right mark, and practical wisdom makes us take the right means. (VI.12, 1144a6)

Judging the right means of achieving the right mark is highly dependent on the details of the situation in which we are immersed, and practical reason will call on its contextual aspect (§3.124). Aristotle makes it clear that such decision making is not a matter of deriving specific instructions for action by deduction from absolutely true general principles, for:

fine and just actions... exhibit much variety and fluctuation... We must be content, then, in speaking of such subjects and with such premisses to indicate the truth roughly and in outline, and in speaking about things which are only for the most part true, and with premisses of the same kind, to reach conclusions that are no better.
(I.3, 1094b14)

Ethical judgement is an art, deriving conclusions that are imprecise and contestable from maxims which are likewise. This is not to claim that there is no ethical truth, that all is relative, because we can address the truth “roughly and in outline.” Rather, it is to acknowledge the fuzziness and fallibility of ethical judgements, the continually open possibility that although we can often distinguish worse from better, we might always be able to do better still.

In short, this survey of Aristotle’s virtue ethics shows that it exhibits the following features on which I shall draw for my own account.

1. there are many morally salient factors (virtues, in Aristotle’s account). These may, in multi-dimensional theories, be in tension with one another, though Aristotle’s doctrine of the unity of the virtues denies this latter feature in his case.
2. the morally salient factors have diverse roots. In Aristotle’s account, genetic factors, social experiences and practical wisdom contribute to morality. Thus there are roles for instinctive, habitual and deliberative inputs into moral actions, rather than solely rational inputs;
3. all applications of the morally salient factors are contextualized;
4. moral judgements cannot be exact, so that they require one and only one outcome;
5. the focus of morality is persons, not acts or principles;
6. we are responsible for, and should care about, our character and the actions that flow from it.

5.21 Moral Development of Children: Instinct and Habituation

Since we are interested in an account of morality that can underpin a program of moral education, however, we need to look more carefully at what our meta-ethical theory has to say about the process by which we come to be moral persons. So far, we have hardly considered Aristotle's views on this question. Above, I identified three roots for morality within his theory: instinctive, habitual and deliberative. These three roots correspond to the three noted at the end of §5.1: the genetic inheritance of humanity, the social situations in which humans are raised, and the influence of reasonableness.

In the first case, humans may have acquired through evolution a genetic disposition to react with a feeling of moral rightness or wrongness to certain events (§2.2). Recent research supports the contention that empathy with others may well have a genetic base, being exhibited from only a few days after birth (Damon 1996). Prior to experience and prior to reasoning, a child has an inbuilt propensity to have moral reactions to at least some situations. Of course, there are other instinctive reactions, not all of which have anything to do with morality. Aristotle concurs with the view that children have inbuilt "proto-virtues" available from birth:

For both children and brutes have the natural disposition to [virtues], but without reason these are evidentially hurtful.... Therefore... in the moral part there are two types, natural virtue and virtue in the strict sense, and of these the latter involves practical wisdom. (VI.13, 1144b7)

But why should it be said that these natural virtues are "hurtful" without reason? The answer must be that instinctive reactions are programmed, in the sense that they occur automatically in response to certain triggers in the environment. The environment in which morality applies - that is, the social environment - is much more complex than the environment in which many instincts play. Situations similar in many non-normative respects may have very different social import to different actors, and hence natural virtues might often lead to inappropriate reactions. Damon quotes the example of the young child who, reacting to another child's crying, leads his own mother over to comfort the other. If the other is crying in the first place due to the presence of strangers, this may in fact make the situation worse. Thus, the hurt that would follow from allowing children to act solely on the natural virtues must be avoided, at first, through controlling children's behaviour and modelling appropriate behaviour. Soon, however, the natural virtues must start to come under the influence of reason, so that childish virtue may be properly situated and applied, even in the absence of an adult.

Let's look a little more closely at the process of controlling and modelling behaviour, which can be considered as important elements of habituation.

Interpretations of Aristotle's concept of habituation and views on moral education have relied heavily on passages such as the following:

a young man is not a proper hearer of lectures on political science; for he is inexperienced in the actions that occur in life, but its discussions start from these and are about these; and, further, since he tends to follow his passions, his study will be vain and unprofitable, because the end aimed at is not knowledge but action. And it makes no difference whether he is young in years or youthful in character; the defect does not depend on time, but on his living and pursuing each successive object, as passion directs. (I.3, 1095a2)

Anyone who is to listen intelligently to lectures about what is noble and just and, generally, about the subjects of political science must have been brought up in good habits. (I.4, 1095b3)

Before argument and teaching can be effective, Aristotle asserts that:

the soul of the student must first have been cultivated by means of habits... The character... must somehow be there already with a kinship to virtue. (X.9, 1179b30)

These passage leaves open two important questions. What degree of formation of character is necessary before teaching can be effective? When is the soul or character sufficiently cultivated? The standard answer seems to be that the character must be fully formed, and that this does not occur until into adulthood. Sherman (1989, 157) comments that "under traditional interpretations... practice is seen primarily as a non-rational training of desires towards appropriate objects." Pritchard (1992), quoting the I.4 passage above (albeit in a different translation), attributes a similar view to Aristotle. Baron (1985, 147) also accepts this view, adding that virtue ethics implies

a form of education [which] would set the subject's gears in a certain way, without his realizing that anything has been set and, of course, without him being in a position to 'reset' the 'gears.' It keeps from the subject both self-understanding and self-direction.

Habituation, on this traditional view, is something that takes place *before* any rational discussion or study of morals can occur. Sherman (op. cit., 194) asserts:

But to postpone inquiry into the ultimate origins of a discipline is obviously not to postpone all inquiry relevant to learning the facts.... Nothing Aristotle says precludes the educational path being marked by stages of inquiry and explanation or there being an explanatory dimension to the acquiring of adequate habits.

Certainly we can agree with Aristotle that experience in the actions that occur in life is a necessary condition for engagement in practical discourse, but this does not fix when such reasoning can be applied. Since even the youngest child capable of speech is already "living and pursuing" objects under the direction of passion, it

seems that at least one of Aristotle's conditions for benefiting from practical reasoning is already met by such a child.

Sherman argues that Aristotle's conception of habituation cannot be merely "non-rational training." In support of this view, she offers a illuminating analysis of the idea of repetition of a process as a way of habituating it. Taking much more simple skills, such as a physical action skill, she notes that even the habituation of such actions cannot be merely a matter of repeating them over and over. This is because the first attempts at an unlearned action are bound to be clumsy and inaccurate. To merely repeat such an action without subjecting it to critique and improving it ensures no progress. Consequently, the learner must make trials which are aimed at an ideal, and this requires awareness of the goal and continued judgements about the closeness of the learner's approach to it. Indeed, Sherman's analysis can be extended here. Every sports coach knows that coaching a raw beginner is easier than coaching someone who has developed bad skill habits. The tyro can be taught properly from the start, whereas the poor skill habits of the experienced player need to be broken and remade. Aristotle makes a similar observation with respect to moral habits.

It is hard, if not impossible, to remove by argument the traits that have long since been incorporated in the character. (X.9, 1179b16)

I conclude, with Sherman, that to allow children to believe that certain actions are good no matter what the circumstances is to lead them away from ethical behaviour. If we are to bring children up to be ethical, we must begin to introduce practical reasoning as soon as they have some habitual responses to build upon. It is easier for reasoning and argument to be involved in moulding the traits of character while they are still in the process of being incorporated, presumably provided that the child is capable of comprehending the reasoning. The discussion of the instillation of sports skills above led us to a similar conclusion.

This conclusion is certainly supported by modern research that indicates that the moral training of young children is more effective if the demand for moral behaviour is accompanied by reasons and discussion (this has been labelled "authoritative parenting"), than if it is imposed in an authoritarian manner (see Damon 1988, Chapter 4, for an excellent summary). Children raised in authoritarian households, through mechanical habituation, lack precisely the initiative and inner responsibility to which Aristotle alluded when he discussed the "man who takes care" in moral matters (III.5, 1114a3, quoted above). We can note that the same research also indicates that reasoning, in the absence of the enforcement of behavioural

expectations (i.e. permissiveness), is also ineffective, thus supporting arguments for the importance of habituation.

Our characters are never fully formed: we are always changing, perhaps becoming more virtuous, perhaps backsliding. If we care about our moral standing, part of the process of becoming more virtuous inherently involves practical reason in the evaluation of the present state of this project, and the decision to act in certain ways so as to reinforce good, or replace bad, habits already formed. Some habits must be in place for this process to work on, and even very young children have already been trained and habituated to some degree. The processes of being lectured on, engaging in argumentation about, and reflecting on, our habituated responses need not, ought not, and cannot wait for a (hypothetical and unachievable) completion of habituation.

The thrust of my argument so far is that, while habituation is clearly an important part of growing up to be moral, we cannot support a characterization of habituation that, like the traditional interpretation of Aristotle, claims the task of applying practical reasoning to the character must wait until virtuous action has been fully habituated. Sherman (1989) in a lengthy discussion of the habituation of character argues that the notion of habituation, as used by Aristotle, is one that inherently involves the use of practical reasoning. Drawing in part on the work of Burnyeat (1980) and Sorabji (1980), she says:

My overall claim is that if full virtue is to meet certain conditions, then this must be reflected in the educational process. This will require a developmental conception of cognitive and affective capacities, as well as a conception of habituation in varying degrees reflective and critical. (159).... We misconstrue Aristotle's notion of action producing character if we isolate the exterior moment of action from the interior cognitive and affective moments which characterize even the beginner's ethical behaviour. (178)

It is worth quoting some of Sherman's arguments for these conclusions at greater length:

The child is not an empty box in which beliefs are instilled, but an individual who has, to a greater or lesser degree, already formed certain construals and judgements, which become adjusted and revised through interaction with an adult. Education is thus a matter of bringing the child to more critical discrimination.... What is required is a shifting of beliefs and perspectives through the guidance of an outside instructor. Such guidance cannot merely be a matter of bringing the child to see this way now, but of providing some sort of continuous and consistent instruction which will allow for the formation of patterns and trends in what the child notices and sees.... Though the educator persuades and exhorts, the goal is not to *manipulate* beliefs and emotions - to influence an outcome here and now - but to prepare the learner for eventually arriving at competent judgements and reactions on his own. Any method which secures rational obedience must at the same time encourage the child's own development... in a way that engages his own critical capacities. What is required is some dialogue and verbal exchange about what one sees and (feels) and should see (and feel); in other words, actual descriptions which articulate a way

of perceiving the situation and which put into play the relevant concepts, considerations and emotions. (171-2)

Throughout this subsection, I have been looking at Aristotle's account of two roots for morality: the instinctive and the habituated. Sherman alerts us here to the interconnections between these two starting points of the child and the role of interaction both with the world and with other persons. Most of all, she highlights the absolute indispensability of reason in the development of both the instinctive and the habituated responses. It is this enhancement of the judgements involved in situated reactions that enables habituation, which might appear on the surface of it to be inherently heteronomous, to play its role in the development of autonomy in the child.

Baier (1985) also melds the genetic (natural responses), the habituated (training in the interdependent situation) and the reasoned (critical element):

If we see morality as essentially control of our natural responses to the mixed risk of evil and chance of good in our interdependent situation, we can see moral response, in the form of training, criticism, and so on, as a response to a natural response... (220). What I have tried to make persuasive is the suggestion that morality begins in a response to a natural response to a situation in which hope and fear are properly inspired by one and the same situation. Morality is a proliferating succession of responses, individual and collective, to those primitive responses to the risky adventure of interdependence... Morality is throughout responsive to already given responses, and its norms are reflective versions of natural responses to the risks and opportunities interdependence involves. If we can see morality this way, I think we can keep that *critical* element which most normative theorists want to stress without making the critical morality one that cannot make constructive critical responses to the real situation. (222-3)

Indeed, in this passage, Baier stresses one of the key advantages of virtue ethics, due to its multi-dimensional character: through its ability to link together emotions and drives, habituated actions and critical reflection, it has a greater potential to explain how it is that moral judgements can coordinate with real, unique situations to lead to moral action. The young child starts moral life through natural reactions to everyday situations that are saturated with social mores. These provide the elements of an enacted morality. Yet, the social interaction in which the child is immersed does more than inculcate moral action. It also inducts the child into the social world of discussion and argument, with the result (as we have seen in §3.23), that the child becomes progressively able to internalize the mechanisms of public dialogue as private thinking. Amongst the topics on which this dialogue and reflection are turned are ethical and moral matters. The child becomes increasingly capable of entering practical discourse (in the Habermasian sense), both in dialogue and in reflection, often assisted by the explicit intervention of more capable reasoners (§4.14). The critical element to which Baier refers enters into the child's ethical life.

This subsection has shown that habituation needs to include engaging children in practical reasoning, through pedagogic action (§4.141), from an early age, if true virtue is ever to be achieved. Yet it has left the issue of the development of practical reasoning, as used by the child, somewhat obscure. Despite Sherman's reference above to the requirement for a developmental conception of cognitive and affective capacities, there has been something of an assumption that practical reason will be there when it is needed, with little explicit treatment of how this might occur. I shall now turn to the question of how practical reasoning develops, and what implications this has for a virtue ethics meta-ethical theory.

5.22 The Development of Practical Reasoning

In the previous subsection, I explored one of the contentious issues in the interpretation of Aristotle: the extent to which he felt that practical reasoning should be withheld until after virtuous habits had been established. It is notable that Aristotle is not subject to the criticism that he ignored development in his meta-ethical theory, but his remarks about it are somewhat lacking in detail, and thus subject to many interpretational controversies. I don't wish to enter into the issue of the proper interpretation of Aristotle on this point, but the controversy does indicate a problem for a multi-dimensional account of morality. If moral education involves both habituation and reasoning, and some measure of the first must be in place before the reasoning can be used, then when ought reasoning to be introduced? What level of practical reasoning ought we to expect children to be able to use, and to begin to internalize, at what ages? These are clearly pressing problems for the teacher of moral education.

I will mark out several positions on these questions to provide starting points for this discussion. At one extreme, we have the view that children ought to be thoroughly habituated to follow what the virtues require before entering into any degree of practical reasoning. The arguments in the previous subsection show that this position is untenable. Some degree of practical reasoning is an integral part of habituation.

At the other end of the spectrum would be the view that children are already capable of engaging in practical reasoning at any age. This position is also untenable: children clearly undergo development in their ability to reason as they grow older. What is not always recognized is the implication: that adults, too, may be at various levels of ability in reasoning. Unless there is some threshold value of adequate and sufficient reasoning that all adults reach upon attaining maturity, we must expect that all judgements of ability to engage in practical reasoning involve placing the level of

reasoning on some scale or scales. In view of the multi-aspectual account of reasonableness I developed in Chapter Three, my view is that we need at least five quasi-independent scales in order to measure reasoning ability.

So the answers to the questions posed in the opening paragraph of this subsection require us to make nuanced and situated judgements in the case of each particular child, or group of children. Clearly, such an answer implies that the particular judgements are empirical matters, to be decided in action. Nevertheless, some philosophical groundwork can assist us by clarifying different levels of practical reasoning.

The most immediate level of practical reasoning will be that which is tied directly to the situation. A parent who prevents a child from hitting another may say “Don’t hit Alan - you can see that he is getting upset.” Here, the parent is making the child do the moral thing, but is also reasoning with the child - offering an immediate and situated reason why the action is not acceptable. Such reasoning commonly (but not necessarily - the parent might just have said “Don’t!”) accompanies coercion in the interests of habituating a moral response. We have seen that research evidence supports the view that such reasoning (in authoritative parenting) is more effective in instilling moral virtues than mere imperatives (authoritarian parenting).

What of the most abstract level? Perhaps it is Aristotle’s “hearer of lectures on political science” (I.3, 1095a2). If Aristotle is the lecturer, then we are talking about very high level reasoning indeed. Given the subtlety and abstraction of such lectures, it is clear that young children will not be capable of handling them.

Between these two levels lie many others. They differ on a number of dimensions: the complexity of the language used; the abstractness of the concepts; the distance from the immediate situation; the difficulty of the logic employed; the degree to which they are initiated by the adult or the child; the degree to which concepts are problematized, and so on.

At this point I want to make a rough distinction between reasoned moral instruction and engagement in philosophical moral discourse. Both of these lie in the multi-dimensional space between obeying authoritarian commands and listening to Aristotle’s lectures. The distinction can only be rough, because the one can almost imperceptibly shade into the other as a discussion builds, and because the character of the discussion can vary along many of the above dimensions quasi-independently. However, we can use a number of these dimensions to help clarify it, as shown in the following table.

Dimension	Reasoned moral instruction	Philosophical moral discourse
1. Abstractness of the concepts	Concepts are tied to concrete exemplars.	Concepts are treated in themselves.
2. Distance from the immediate situation	Discussion tied to immediate situation, or a similar class of situations.	Discussion refers to generalized situations.
3. Pattern of initiation of discussion episodes	Instructor wields control over topics of discussion.	Control over topics shared amongst group.
4. Degree of conceptual clarification	Concepts taken for granted, to be used in guiding action.	Concepts themselves become topic of conversation.

I shall illustrate this distinction by reference to a suite of related examples, clustered around the moral concept of fairness. A typical example of reasoned moral instruction might be when a teacher sees a child taking most of the coloured pencils from the table's supply. The discussion of fairness that follows is likely to revolve around the issue of the pencils and related examples ("How would you feel if there were five cakes on the table and Greg took four?"). Here, fairness is exemplified by concrete distribution problems, and tied to the issues of the pencils present and similar cases. The teacher decides how the conversation proceeds and what issues are introduced. The teacher has the idea of fair, and wishes the child to acquire it. Reasoning is used, but to achieve the aim of instilling good sharing habits among the children at the table.

In a philosophical moral discourse involving the same children, the children might explore what the word fair means - does it mean, for example, equal distribution? The discussion may not refer to any particular situation that involves the children, but range across a number of imaginary examples, often at a quite generalized level (e.g. the distribution of goods rather than pencils). Each participant in the discussion can introduce new questions or assertions, and many of the words used (e.g. equal, deserve) might themselves be problematized. Reasoning here goes beyond merely instilling good habits (though it may still serve that purpose), and embarks on the journey of finding out what good habits are, and what makes them good.

It seems clear that habituation into the virtues must, as I have argued, involve reasoned moral instruction, at any age, provided the reasoning is appropriate to the abilities of the child(ren) involved. It also seems clear that philosophical moral discourse requires a greater ability to reason (for the same level of conceptual complexity), and hence must in some sense be later than critical habituation. But we

still have not addressed the question of the development of the reasoning capacity that both require.

Sherman (1989, 190-7) argues at some length that critical habituation is an integral part of becoming morally good and, further, that not just reasoned moral instruction, but also a certain amount of what I have called philosophical moral discourse is essential to critical habituation. The young must investigate how to construe situations, how to proceed, how to distinguish morally salient factors, how to live. All this must occur before the young are equipped to attend Aristotle's lectures. She concludes:

Thus if we are to become critical listeners of discourse on ethics, we must first be trained to become critical inquirers. Ethical education must, to some extent, include this training.

While I concur with this judgement, I wish to go further. What is left unclear in Sherman's account is the process of training the young to become critical inquirers. I have been developing an account of moral development which gives a central place to the engagement of children in moral philosophizing, at the level appropriate to their development. This, I will argue, is an important element in training them "to become critical inquirers" in the context of ethical education. Thus, I will be arguing that Sherman has oversimplified: becoming critical inquirers in ethical matters can be achieved through (amongst other things) becoming critical listeners - and later, speakers - in discourses on ethics. In other words, becoming a critical inquirer is achieved (best, possibly only) by engaging in critical inquiry (Vygotsky 1962; 1981, see §3.23). In the early stages, engaging in critical inquiry just *is* listening in to inquiries carried on around us. With growing experience, the ability to participate in the inquiry is gradually internalized and implemented (see the account of the acquisition of competency in playing "peek-a-boo" in Bruner 1983 as an exemplar of this process). I shall return to this in §5.232 below.

Although I have framed this as a Vygotskian claim, there is a way of reframing it that is clearly Aristotelian. His analysis of the development of moral virtues is applicable also to the intellectual virtues - in particular, practical wisdom - in a way that undermines any interpretation of Aristotle which claims that the young cannot benefit from philosophy. The engagement in moral philosophy is the use of practical wisdom. Hutchinson (1995, 207) gives a useful summary of Aristotle's account of practical wisdom:

All in all, practical wisdom is an appreciation of what is good and bad for us at the highest level, together with a correct apprehension of the facts of experience, together with the skill to make correct inferences about how to apply our general moral knowledge to our particular situation, and to do so quickly and reliably.

Practical wisdom, we can see, contains three parts: moral appreciation (or an appreciation of the moral virtues); apprehension of facts; and practical reasoning. Now, practical wisdom is one of the five intellectual virtues (I.13). The moral virtues are, as we have seen, based on habituation - a critical habituation. Is the same true of the intellectual virtues - in particular, is it true of practical reason? I have advanced the claim that it is, using a somewhat different terminology. The way that we become able to reason, involves an essential element of habituation, on the sort of Vygotskian account I have been urging. We become able to make certain intellectual moves because we have been engaged, repeatedly, in conversations where those moves have been made. Reid (1872, 641) makes a similar point:

Our judgment of things is ripened, not by time only, but chiefly by being exercised about things of the same, or of a similar kind.... I am very apt to think, that, if a man could be reared from infancy, without any society of his fellow-creatures, he would hardly ever shew any sign of moral judgment, or of the power of reasoning.

In other words, moral judgment and reasoning is something that needs to be learned from others, and it must be practiced in the company of others, who are already better at it than are we, and in contexts sufficiently like those in which we are to subsequently use it. To return again to Sherman's quote, to become critical inquirers involves firstly being critical listeners (and critical speakers).

We might paraphrase Aristotle here: being capable of practical reasoning is not merely being capable of making intellectual moves. It is a matter of accessing the right intellectual moves at the right time, with reference to the right subject matter, in interaction with the right people, with the right motive, and in the right way. The habits of intellect need to be reflexively turned upon themselves, if we are to reach practical wisdom. For the same point applies to habituated reasoning as to habituated virtue - if we allow the habituation of poor or sloppy reasoning, then it becomes very difficult to set it right later.

So we cannot treat practical reasoning as a given - something that is merely applied to moral appreciation and the apprehension of facts. Practical reason is itself the result of much habituation and reflection. Practical reason ought to develop hand-in-hand with habituated virtuous action, if either is going to be developed to anything like the desirable extent, for both are the product of practice and critique.

5.23 Children and Moral Philosophy

I do not seek to establish that the young ought to be sent along to high level lectures in moral philosophy, akin to Aristotle's, nor that the discourses on moral philosophy that they engage in *have* to be formally convened (although I will argue that it is a

good idea to do so). At least some of the ordinary, everyday discussion on moral matters in which authoritative parents and carers engage with children meets the criteria advanced above for philosophical moral discourse, albeit in a fragmentary and loosely structured way. My claim is rather that some involvement in philosophical moral discourse of some type is an integral part of effective moral development.

Yet, as we saw in §5.21, a widespread interpretation of Aristotle holds that he claims that none but the already experienced, morally habituated person can get anything out of *any* engagement in moral philosophy, at no matter what level. Such an interpretation depends on a conflation of Aristotle's lectures with moral philosophy in general. For example, even such a sympathetic commentator as Sherman conflates "explanation... of the sort that Aristotle's lectures on ethics can yield" with "some theoretical and general account of the substantive and formal features of good living" (Sherman 1989, 194). The latter can be addressed at a level considerably below that at which Aristotle addresses it - a level that lacks his rigour and breadth, but which is still recognisably philosophical and well within the reach of the young.

Nussbaum (1994, 97), another commentator sympathetic to the place of practical reasoning in habituation, makes a similar, but longer, slide in worrying about "a tension in Aristotle's position" between his partially cognitive account of emotion and this requirement to fully form the moral emotions before studying moral philosophy. She says:

On the one hand, he describes the emotions as closely bound up with judgments, and therefore capable of being modified by the modification of judgment. This picture implies not only that emotions can play a role in rational deliberation, but also that they can be changed as beliefs of all sorts can be changed, *by* deliberation and argument. On the other hand... he makes a sharp distinction between character training and the philosophical study of ethics, on the grounds that the emotions need to be balanced *before* the student can get anything much out of his philosophical arguments. Why, taking the view of emotions that he does, does he appear to insist on a separation between character training and philosophy? Why can't philosophical argument shape character?

"His [Aristotle's] philosophical arguments"² turn in the subsequent two questions into "philosophy" and "philosophical argument" in general (and note that these latter activities can be much less formal than Sherman's "theoretical and general account of the substantive and formal features"). What if we were to compare this to a modern parallel? If, say, Steven Hawking were quoted as saying that no-one could benefit from his lectures without having become thoroughly habituated to tensor

² Presumably as delivered in a lecture, not as studied at leisure.

calculus, would we conclude from this that children ought not to be exposed to mathematical inquiry? I venture to think not.

Nussbaum's tension dissolves if we take Aristotle to be talking only about terse, dense, verbal delivery of a tightly argued system, and not about all philosophical inquiry. Indeed, it is difficult to see how he could be talking about the latter, since much general discussion has a philosophical aspect; there can be no sharp line drawn between philosophy and other types of inquiry.

But let's leave aside the exegetical question of whether restricting the scope of Aristotle's reference to encompass only formal, high-level lectures is legitimate or not. His remarks have often been taken to be a claim that children (not to mention women and young men) are incapable of either doing moral philosophy or benefiting from it. And whether they draw their view from Aristotle or not, there are those who still hold this position.

There can be two sorts of attack on this view. One is empirical, and can be mounted by pointing to examples of children doing moral philosophy, or by attempting to discuss moral philosophical issues with groups of children and assessing whether they do, in fact engage in something that is recognizably moral philosophy. I do not intend to take this approach here. Others have, and have compiled an impressive amount of data to show that children can and do philosophize, from as young as four years old (see, for example, Matthews 1980; Matthews 1994; Pritchard 1996).

The second approach is to attack philosophically the view that children are unable to engage in moral philosophy. Such an attack will have two prongs, both of which have been foreshadowed above. Firstly, I will argue that philosophy is not such a radically different undertaking from general discourse about ethical matters that a sharp wedge can be driven between the two. Secondly, I will draw on the understanding gained above concerning the development of reasoning, particularly practical reasoning, to argue that the recognition of the philosophical element in moral development warrants more explicit study of moral philosophy in order to assist children to become ethical agents.

5.231 Philosophy and Everyday Discourse

Philosophers who are considering the place of moral philosophy in the world quite naturally construe philosophy as that in which they are engaged. Philosophy is reading the greats (or, in journals, the not-so-greats); it is attending learned seminars; it is writing learned and technically difficult articles and so on. To claim that the fact that children, or even ordinary people, are not capable of participating at this

level is no argument that they are not capable of philosophizing at any level. My fourteen year old son's inability to do integral calculus does not rule out his handling quadratic equations. Indeed, to be able to eventually handle integral calculus, it is imperative that he study lower level mathematics for a long time, and that this was grounded in informal, everyday mathematical awareness long before he started to attend school.

For the ordinary member of the public, however, philosophy is something quite different. It is something like an approach to things, an underlying way of looking at the world, a listing of important guides to life. It can be talked about in ordinary language on ordinary occasions - around a coffee table, or over a beer - whenever we slow down a little and think a bit more reflectively. We pass on our philosophy of life to others whenever we muse a little and talk about what is important to us in the big picture. Often, we do this when we are talking to our children, or to the young in general.

This everyday conception of philosophy is quite different from Aristotle's lectures in many ways. It is much more diffuse, less rigorous, more poorly connected and systematic. Yet it also shares some important features. It is about "the big questions," included amongst which is the question "How ought I to live my life?" Answers to this question are, in the main, philosophical answers, whether they arise in the context of a tightly argued lecture series, or in epigrammatic utterances. Of course, the *quality* of the answers, as philosophy, varies in the two cases (though we need to be aware that the quality as advice for life might not covary with the philosophical quality).

This characterization of philosophy as attempts to ask and to answer the "big questions" (and I don't wish to attempt to spell out just what makes a question "big," or to list them) makes it clear that philosophy is something that most - probably all - people engage in at various times. In particular, any attempt to bring up children to be moral that extends beyond mere mechanical habituation, that attempts to give reasons for actions and attitudes, is to that extent philosophical. And it is not merely reasoned moral instruction: the child who questions and explores these issues is in the realm of philosophical moral discourse. As both Sherman and I have argued, such an extension is essential to any effective habituation in the ethical, and so all but a small minority of severely ethically disadvantaged children do engage in at least a low level of philosophical inquiry in the course of growing up and acquiring an ethical outlook.

If we accept that there is an inherently philosophical element in ethical education, even if it is quite unfocused, then the question arises as to the degree to which it is useful to make this element more focused in the ethical upbringing of children. At this point, it is useful to recall the spectrum of “moral philosophicality,” extending from a few wise saws offered with little opportunity to explore them, through to Aristotle’s lectures.³ Aristotle, clearly, thought that a course in highly abstract moral theory was of no use, and it is hard to disagree with this view. Nor, in the light of the arguments presented above about the need for the development of nuanced judgement in the building of virtue, can we accept that the odd “philosophical advice” without conceptual exploration will suffice. The answer lies somewhere in the middle - but exactly where?

5.232 Children, Moral Philosophy and Practical Reasoning

My answer to this question brings me to the second strand of the present discussion: the development of practical reason (that is, reasonableness as applied to moral situations) that must parallel the development of good moral habits. We have established that habituation cannot be a blind, repetitive process. It must contain a strand of critical evaluation. Now, Aristotle is quite right that a failure to habituate the young into good moral habits makes the attainment of practical wisdom much more difficult in adulthood (e.g. X.9, 1179b30 and 1180a14, quoted above). Since the same applies to the attainment of practical reasoning - it too must be habituated, and this habituation involves critical evaluation - then sound ethical education must involve both the insistence that children act morally and the insistence that they learn how to judge whether their actions are moral or not. Being able to do the former requires the latter.

Let’s analyse further what the latter means, and how to attain it. Being able to judge whether an action is morally good or not involves having at least the following:

1. a model of the morally correct action presented to one through speech and/or observation. This model may have been presented in the past, or it may be being presented at the time of acting;
2. a concept of the target morally correct action in mind, derived from that presentation, and other previous presentations;

³ Of course, a single dimensional spectrum is too simple. “Philosophicality”, as we saw, has many dimensions. Further, it could be argued that the ‘wise saw’ end of the spectrum has lost too many philosophical features to be properly called philosophy - it is, rather, reasoned moral instruction. Common usage, on the other hand, seems to disagree. However, these are not topics that need deeper exploration here.

3. an accurate picture of how one is presently performing the action, so as to be able to compare it to the model;
4. a concept of the salient features of one's action, and how they relate to the morality of the action;
5. a grasp of the external features of the situation in which the action arises, including an appreciation of the others involved;
6. an understanding of oneself, so as to be able to relate the action to the moral goals and aspirations that one has, for these also help form the situation in which one is immersed;
7. the necessary intellectual tools to be able to draw conclusions from the information available.

Only point 1 is captured by a version of habituation that sees it as externally imposed repetitive action; i.e., mechanical habituation. It is the only one that *must* be presented by another, presumably one who has a greater grasp of morally appropriate action. Let's call this person the moral trainer. Each of the other six points may be present in one of three ways: in the moral trainer; in the moral trainee; or in the jointly constructed interpersonal space that includes the moral trainer and the moral trainee.

A mechanical model of habituation assumes that all seven points are the province of the moral trainer. Leaving aside considerations of how communication can be established at all - considerations which would only strengthen the analysis to be presented here - then mechanical habituation asserts that the trainer not only provides a model for action, but also carries out the assessment of all the other points, and then communicates the conclusions drawn from them to the trainee, along with the information on how the action is to be modified. Constant repetition results in the inculcation of virtuous habits, until the stage is reached where the trainee is ripe for moral wisdom. This account leaves the process of transformation from unthinking trainee to moral wisdom candidate entirely unexplained.

The second way in which the last six points are present - as features of the trainee - is clearly the endpoint at which the process of ethical education is aimed. If the trainee is fully capable of instantiating all seven points, then there is no longer any need for the trainer.

It is the third possibility that provides a way of transforming the first situation into the second. The first situation, where the moral trainer has responsibility for all seven points, may not be an accurate description of the process of habituation throughout, but it is certainly close to the truth for the very young child, just starting out on the route towards moral wisdom and independence. Parents of very young children do closely monitor not just the actions of the child, but also the various features of the situation and even take it upon themselves to decide the interests that the child ought to have.

Yet this situation rapidly changes. The parents negotiate with the child, drawing attention to salient features, talking about the concepts involved, asking the child what they desire and so on. The child reacts. In this intersubjective space, a joint understanding of the moral import of the actions is constructed, and the child begins to build moral, conceptual and intellectual competencies. These do not appear overnight: the parents and other interlocutors of the child continue to provide scaffolding, gradually passing over competence. The child internalizes the competencies. This is, of course, the Vygotskian picture outlined in §3.23. Bruner (1983) maps the process, albeit in the context of learning the game of peek-a-boo rather than learning moral actions. Only such an account of moral learning is adequate to explaining how it is that habituation, understood richly, plays an important role in moral education.

There remains one more task, however, to complete this section. That is to substantiate the claim that the sort of interpersonal negotiation of meaning and competency outlined is philosophical, and that a more explicit recognition of this is important in ethical education. I shall approach the first part of this task by considering a number of the points in the outline of the judgement of moral action presented above, and show that they are, at root, philosophical points.

Having a concept of the target moral action in mind - the second point - involves the conceptual clarification of key moral concepts. Conceptual analysis is clearly a core philosophical pursuit. Yet the clarification of key moral concepts need not take place through explicit conceptual analysis. Many people become able to use concepts adequately through everyday interactions with others. Is this philosophy? We can take a number of attitudes to this question. We can deny that it is philosophy in any sense of the word, because it is something imbibed, something learned through an unconscious induction from examples. Or we can call it a simple kind of philosophy - folk philosophy, perhaps - noting that we do at least sometimes turn our attention informally to the meaning of such concepts. Either way, it ought

to be clear that to make the process of exploring concepts more explicit would be to make more likely the improvement of our conceptual grasp of key moral concepts.

The third point - being aware of the way in which one is performing the action - involves a metacognitive awareness of any or all of the embodied, the committed and the contextualized aspects of reasonableness. Although these may not be specifically philosophical inquiries, we shall see in §10.223 that such narrative and psychological inquiry can play an important part in the community of ethical inquiry.

The description of any action can be infinitely expanded, and some features of any action one takes will be irrelevant to its moral import. Thus it is necessary to have a concept of the salient features of one's action, and how they relate to morality - the fourth point. For example, if I am repaying a debt I owe you, then it would generally be morally irrelevant whether I give you the money with my left or right hand (at least in the West), or what mix of denominations of coins or notes I give you. The latter becomes morally relevant, however, if I deliberately choose to give you a large number of small denomination coins so as to inconvenience you. Even this simple example shows that it is not necessarily straightforward to decide what features of a situation have moral relevance. A feel for moral relevance can be attained experientially, but an explicit, philosophical exploration of the situated criteria for moral relevance will assist in moral judgement of the features of actions so as to know how to adapt them to more closely match the target virtue.

A similar consideration applies to the features of a moral situation that are external to the self. It is necessary to have a grasp of the external features of the situation within which one acts, and this must include an understanding of others. There are similarities between this fifth point and point 4, as external situations are also complex, and the relevance of their features to moral matters may also not be clear. But we need to notice that a further complication is at work here. An understanding of others requires some idea of how it is that they construe the world, and what matters to them. Through a philosophical exploration of moral matters, others begin to express their views and thus start to reveal themselves further. By means of our engagement with these views, we may begin to merge horizons and create an intersubjective space in which the morally charged actions of others make more sense to us.

The creation of intersubjectivity referred above is a mutual, interactive process. In coming to understand others, we also come to understand ourselves better. In mutual dialogue, we can come to know ourselves as we come to know others.

Exploration of what *eudaimonia* means to us, helps us to form ourselves. It is not a given, that we merely need to explore to understand, but a self-conception that is both revealed and shaped by philosophical reflection with others.

The final point on my list refers to the construction of reasonableness which, as we saw in §3.23, takes place in social interaction. The tools of reasoning, practical and theoretical, are one of the central concerns of philosophy (Lipman 1991, 27), and the engagement in philosophical inquiry is the high road to the improvement of thinking. Lipman's book is a justification of his claim that "the capacity of philosophy, when properly reconstructed and properly taught, to bring about higher-order thinking in education [is] significantly greater than the capacity of any other approach" (3).

5.3 Summary

In this chapter, I firstly mapped out some of the features that an adequate base for moral education would need to have. In order to account for the complexity of moral development, we saw, we need a multi-dimensional meta-ethical theory. The theory presented by Aristotle in the *Nicomachean Ethics* - that is virtue ethics - showed promise of having many of the features required, provided that careful attention is paid to developmental matters. The survey of Aristotle's virtue ethics forming the second section of this chapter has led us to a number of conclusions:

1. Aristotle's virtue ethics, being a multi-dimensional moral account, allows that morality is an "inexact science," so that there is no strict algorithm for calculating the right and proper thing to do in all circumstances. Rather, the virtuous person weighs up a number of competing ethical demands and, taking full account of the particulars of the circumstances, attempts to find the ethical balance. Some of these aspects of morality are those mapped by other, Millian meta-ethical theories (such as Kant's Categorical Imperative based theory, or Mill's own Utilitarianism) which purport to be the single moral principle. Each of these insights, however, can find a substantial place in a multi-dimensional virtue ethics, as the virtues of justice, benevolence and so on.
2. Virtue ethics finds a place for the genetic elements of morality, as it allows that we are prepared by nature to be capable of virtue, through the natural virtues. It is thus compatible with some of the insights of evolutionary ethics.
3. The central role of socialization in the construction of virtuous persons is recognized by virtue ethics, especially once we have become clear about

the intertwined roles of habituation to virtue and habituation to practical reason, which must develop hand-in-glove. So, to bring up a child to be moral necessarily includes both the giving of direction as to what actions ought to be taken, and reasoning about those actions. Neither is sufficient on its own.

4. Close attention to the development of morality in childhood has led us to conclude that we must give a more important role to burgeoning practical reason (reasonableness), and its development through (proto) philosophical inquiry.

There are also a number of lacunae in the Aristotelian meta-ethical theory. I have commented in a number of places that the virtue ethics account has left out any detailed consideration of the place of dialogue in moral education, as emphasized by point 4 above. Indeed, the neglect of ethical dialogue goes deeper than that. When I outlined the seven capacities necessary to being able to judge the moral worth of an action in §5.232 above, the first assumes that the moral trainer already has a clear idea of the morally correct action. In a monocultural society, such as Aristotle's, when moral correctness is less problematic than in a pluralistic society like our own, it is perhaps less surprising that Aristotle's theory pays little attention to public debate concerning normative correctness.

For these reasons, there is a need to supplement virtue ethics with an account of the place of dialogical reflection in ethics. There is a meta-ethical theory that does place explicitly linguistic reflection at the core of ethics: Habermas' discourse ethics. In §6.1, I shall firstly outline the theory of discourse ethics, especially as presented in *Moral Consciousness and Communicative Action* (Habermas 1990) and *Justification and Application* (Habermas 1993). Then in §6.2, I shall critically examine the theory, drawing in part on recent feminist, neo-Aristotelian and communitarian commentaries, but also on the accounts of reasonableness and autonomy developed in Chapters Three and Four. Discourse ethics is an explicitly Kantian meta-ethical theory, but in the light of these commentaries, I shall show in §7.1 that it can be cut loose from its Kantian roots without loss of the insights it affords. Finally in §7.2, I will turn to the development of a positive account of the place of discourse in a virtue theory based ethical education. This will take account of the sound core of discourse ethics, the weaknesses that have been exposed in it, and the needs of children, who are not yet ethical persons in a fully fledged sense, but who are working towards it. By the end of Part II, then, I will have advanced a multi-dimensional meta-ethical theory which is based on virtue ethics and incorporates the important insights of discourse ethics.

In Part III, I shall draw on the concepts of reasonableness and communicative autonomy from Part I, and this meta-ethical account from Part II. These will then be applied to the moral education classroom, and I will develop in detail an account of the place of moral philosophical discourse and reflection in moral education.

Chapter Six: Discourse Ethics

In this chapter, I will explore Discourse Ethics, which Habermas develops as a complete meta-ethical system. I will claim that it fails to provide a full account of morality for five reasons. The first is that Habermas' plausible claim that communication requires a presupposition that all taking part are aiming at consensus (in matters to do with the right) does not establish that such a consensus is inevitable. The second, related to the first, is that Habermas needs to be able to distinguish cleanly matters of the right from matters of the good, but this distinction cannot be drawn sharply. The third problem is that Habermas assumes too easily that the actual consensus reached in real discourse tracks fairly closely the consensus that would be reached in an ideal discourse.¹ Fourthly, Habermas must assume that every participant in the ideal discourse would have access to an identical ideal rationality, and I will argue that we have no guarantee that this is so. Finally, Habermas fails to establish his claim that discourse ethics is based on a purely formal basis, and does not need to rely on some pre-existing substantive norms.

Nevertheless, in the course of this chapter and the next, I will argue that discourse ethics, when properly reworked, has much to offer as a supplement to the virtue ethics account outlined in the previous chapter. In a number of ways, it dovetails neatly with, and provides a richer account for, virtue ethics. In particular, it clarifies the role of socialization in moral development, explains why respect needs to be taken as a base value for morality, provides a yardstick for assessing both moral progress and justice, augments the Aristotelian account of the place of reasonableness in morality, and reinforces, in moving from the monological to the dialogical self, why we must take development seriously in any meta-ethical account.

By the end of Chapter Seven, I will have presented the target that we can aim at in developing moral persons through moral education: a target based in virtue ethics, but clarified through insights from discourse ethics.

¹ Any meta-ethics which does not posit directly perceivable, or revealed, moral truths faces some sort of analogous charge, so this is not in itself a fatal weakness. I shall return to this in §6.22.

6.1 Habermas and Discourse Ethics

Discourse Ethics, as developed by Jürgen Habermas from work by Apel (e.g. 1980) and Peters (1966), is a decidedly Kantian enterprise. Habermas (1990, 65-6 - subsequent citations are to this work unless otherwise indicated) bases discourse ethics on two principles. Firstly, every valid norm has to fulfil the principle of universalization that:

(U) *All affected can accept the consequences and the side effects its general observance can be anticipated to have for the satisfaction of everyone's interests (and these consequences are preferred to those of known alternative possibilities for regulation).*

Secondly, there is the principle of discourse ethics that asserts:

(D): *Only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity as participants in a practice discourse.*

(U) gives us the condition for accepting any particular posited norm as valid: it must be universalizable in a specific and technical sense. The term “universalization” refers to a situation in which everyone who could possibly be touched in any way by the application of the norm, has agreed to accept the outcomes that would follow from the adoption of the norm. They must have forecast what the effects of the norm would be, if everyone followed it, and they must all be happy with the forecast outcomes, in the light of what each of them wants out of life. The parenthetical comment, which does not appear in alternative statements of (U) (e.g. p 93), implies that all the alternatives to the posited norm have been similarly investigated. (U) clearly has an ideal, counterfactual character, for such investigations could never take place in actuality.

(D) provides the procedure (p 103) for deciding which particular norms do, in fact, constitute morality. Morally right action consists in following this set of norms, and we have seen from (U) that the set must be decided by seeking the approval of everyone on whom they might have effects. (D) says that this approval can only be given in a particular way - when the approval has been negotiated by the members of a practical discourse (and anyone who might be affected has to be one of the members). In other words, all the negotiators must withdraw from action for the moment, and concentrate on redeeming the validity of contested norms (proposed and questioned within the lifeworld), solely through observing the force of the better argument (103-4). Again, as particularly emphasized by the parenthetical comment in the definition, this must be seen as an ideal, counterfactual situation. I discussed the conditions of such ideal discourse in § 4.123 and §4.124. There is an assumption here that one

particular argument will be *the* better argument, and that all will be able to recognize this. I shall explore this notion of consensus at further length in §6.21.

One point should be noted here. Habermas is quite clear that this procedure does not allow us to create norms, but merely to redeem them:

Basic norms of law and morality fall outside the jurisdiction of moral theory; they must be viewed as substantive principles to be justified in practical discourses. (86)

Practical discourse is not a procedure for generating justified norms but a procedure for testing the validity of norms that are being proposed and hypothetically considered for adoption.... In its openness, practical discourse is dependent upon contingent content being fed into it from outside. (103)

Norms themselves arise within the lifeworld; they form part of the taken-for-granted background to strategic and communicative action. When any participant calls a norm into question, then practical discourse is the procedure to use to try to redeem the norm. If the better argument goes against that norm, then it can be replaced with the one supported by that argument.

In that it is based on these two principles, discourse ethics is a Millian theory (in the sense developed in §5.1 above). It purports to tell us, in general terms, how we can find moral rightness. It seeks to ground morality, though we must note that the type of grounding being talked about here is not a foundationalist enterprise - at least according to Habermas. Some of his critics do not agree, and I shall address this in §6.24.

Habermas emphasizes one vital respect in which discourse ethics is very different from Kantian ethics. Discourse ethics is a dialogical and not a monological theory, a position that he explicitly contrasts with that of Kant and such neo-Kantians as John Rawls. Monological theories hold that any one person acting alone, provided he or she is a sufficiently good thinker, is able to derive substantive ethical norms by the use of an appropriate solitary and reflective method. A dialogical theory denies this. It will hold that the process of arriving at any substantive ethical correctness must, at least in the ideal, involve discussion between all those affected.

Habermas explains this as follows:

I have formulated (U) in a way which precludes a monological application of the principle. First, (U) regulates only argumentation among a plurality of participants; second, it suggests the perspective of real-life argumentation, in which all affected are admitted as participants. In this respect my universalization principle differs from the one John Rawls proposes.

Rawls wants to ensure impartial consideration of all affected interests by putting the moral judge into a fictitious "original position," where differences of power are eliminated, equal freedoms for all are guaranteed, and the individual is left in a

position of ignorance with regard to the position he might occupy in a future social order. Like Kant, Rawls operationalizes the standpoint of impartiality in such a way that every individual can undertake to justify basic norms on his own. The same holds for the moral philosopher himself. It is only logical, therefore, that Rawls views the substantive parts of his study (e.g. the principle of average utility), not as the *contribution* of a participant in argumentation to a process of discursive will formation regarding the basic institutions of late capitalist society, but as the outcome of a "theory of justice," which he as an expert is qualified to construct.

If we keep in mind the action-coordinating function that normative validity claims play in the communicative practice of everyday life, we see why the problems to be resolved in moral argumentation cannot be handled monologically but require a cooperative effort.... Only an intersubjective process of reaching understanding can produce an agreement that is reflexive in nature; only it can give the participants the knowledge that they have collectively become convinced of something.

From this viewpoint, the categorical imperative needs to be reformulated as follows: "Rather than ascribing as valid to all others any maxim that I can will to be a universal law, I must submit my maxim to all others for purposes of discursively testing its claim to universality. The emphasis shifts from what each can will without contradiction to be a general law, to what all can will in agreement to be a universal norm [quoting McCarthy]." This version of the universality principle does in fact entail the idea of a cooperative process of argumentation. For one thing, nothing better prevents others from perspectively distorting one's own interests than actual participation. It is in this pragmatic sense that the individual is the last court of appeal for judging what is in his best interest. On the other hand, the descriptive terms in which each individual perceives his interests must be open to criticism by others. Needs and wants are interpreted in the light of cultural values. Since cultural values are always components of intersubjectively shared traditions, the revision of the values used to interpret needs and wants cannot be a matter for individuals to handle monologically. (66-8)

Here, Habermas highlights his view that universalization on a Kantian monological model is in fact a process of impoverishment. The Kantian or Rawlsian moral legislator strips away all inclination, all knowledge of self as different from others. This is a least common denominator picture of the subject of moral norms.

In discourse ethics, however, each person represents him or herself. Each brings their own concerns, likes, dislikes, prejudices, convictions and so on to the conversation. If in the end many of these are set aside as irrelevant to the normative judgement in hand, then this is because each has been convinced of the need to do so. So practical judgement starts, not with a process of impoverishment, but with a process of enrichment. What is set aside in universalizing is not what is decided, by one person in advance of even entering into the making of the practical judgement, to be irrelevant, but what is mutually decided, after wide consultation in which all have their say, to be irrelevant. There is no chance that morally relevant features will be accidentally left out because of the systematic blindness of the moral legislator.

Habermas advances another reason (at the end of the quotation) as to why we cannot allow one person, even an idealized one, to make our moral judgements. The understanding of morally charged situations, he asserts, involves interpretation of the

culture within which they are found, which in turn interacts with the values expressed by, and the needs and wants of, the individuals involved. Each of these - culture, individual interests, intersubjective values - is recursively connected to the others, in ways that cannot be established by solitary contemplation, but need to be clarified collaboratively.

Kant and his successors have commonly faced the charge that moral legislators in the “original position” make exclusions of this sort. Feminist commentators in particular have alleged that Kant wrongly left gender out of consideration, with distorting effects on many of the substantive moral principles he derived, and that by further leaving out considerations of interpersonal attachment, Kant’s ethical system is flawed at a deeper, more fundamental level, in that it overemphasizes justice, and neglects care. For these reasons, Habermas’ account of meta-ethics has received a welcome from feminist writers (Benhabib 1992; Meehan 1995), though not without reservations. In §6.2, I shall consider some of these reservations.

We can see, from the second paragraph of the passage quoted above, that Habermas holds that no substantive moral conclusions can be asserted by an individual as *the* valid moral conclusions. At best, all that an individual can do is to submit a proposed norm to the discourse. He criticises Rawls for purporting to tell us of the correct conclusions of substantive moral argument from a monological position, when all that he can validly aspire to is to join the discourse. We cannot, Habermas avers, cast ourselves as moral experts who can monologically answer any moral questions.

On first glance, this seems to catch Habermas in a contradiction. Surely in advancing the principles (U) and (D), he is doing precisely what he has just accused Rawls of doing? Why is it that these principles are not themselves moral claims that need to be submitted to practical discourse? To understand why Habermas thinks that his principles escape such a charge, we need to turn to the ideas of transcendental pragmatic justifications and performative contradiction.

Thus the necessary justification for the proposed moral principle could take the following form: every argumentation, regardless of the context in which it occurs, rests on pragmatic presuppositions from whose propositional content the principle of universalism (U) can be derived. (82)

Habermas claims that (U) is a formal, and not substantive, moral principle. As the previous quotation asserts, (U) rests on those necessary assumptions that we cannot escape if we engage with each other in the attempt to redeem moral principles at all. I investigated what Habermas means by argumentation, or discourse, in §4.123. There we saw that Habermas claims that argumentation is an attempt to decide a question by recourse to reasons alone. If this is to be able to happen, then the following rules

(Habermas quotes their formulation by R. Alexy) must be always already assumed by those entering argumentation:

(3.1) Every subject with the competence to speak and act is allowed to take part in the discourse.

(3.2) a. Everyone is allowed to question any assertion whatever.

b. Everyone is allowed to introduce any assertion whatever into the discourse.

c. Everyone is allowed to express his attitudes, desires and needs.

(3.3) No speaker may be prevented, by internal or external coercion, from exercising his rights as laid down in (3.1) and (3.2)....

If these considerations are to amount to more than a definition favoring an ideal form of communication and thus prejudging everything else, we must show that these rules of discourse are not mere *conventions*; rather they are inescapable presuppositions. (89)

In the last sentence, Habermas clearly recognizes that the status of these rules is crucial. For his account to succeed, it must be shown that the three rules are not just a series of recommendations for improving discourse, but our best representations of ideal requirements that every participant must tacitly understand and accept in order that argumentation could even be possible. The actual form of words used to express these representations, he says, is a rational reconstruction of these presuppositions, and could be inaccurate. In order to establish the inescapable nature of these presuppositions, Habermas turns to the notion of a performative contradiction.

Assertions are acts. So it is possible to perform an act by making a statement in which the assertion and the act performed are mutually contradictory. Consider the statement "Don't listen to me!" The act of uttering this statement is a claim to be listened to, yet the content contradicts this. Habermas claims that in entering argumentation - making any argumentative moves at all - we are always already acting in a way which contradicts any assertion that coercion, exclusion or limitation is acceptable. The rules 3.1-3 must be tacitly accepted in order to enter argument or practical discourse at all (including asserting the rejection of those rules), for in:

practical discourse we always already make use of substantive normative rules of argumentation. It is *these rules* alone that transcendental pragmatics is in a position to derive. (86)

To understand the force of this, we need to remind ourselves of the special sense in which Habermas uses "argumentation": as the attempt to redeem claims by the force of the better argument only. It might seem that the anti-Habermasian skeptic merely has to point out that real arguments and discourses do not rely *solely* on the force of the better argument - that inevitably they involve coercive aspects as well. That such

dialogues take place, Habermas does not deny (though he does withhold the labels “argumentation” and “discourse” from such dialogues). In a long discussion (pp 98-102), he considers what the implications of this retreat from argumentation to different modes of dialogue are. He firstly claims that skeptics must refuse to enter into any form of argumentation at all (pure or not), if they wish to escape the binding force of the presuppositions. That is, the skeptic cannot appeal to better reasons at all, in the attempt to establish any norm, without accepting the presuppositions that underlie such an attempt. Further, because argumentation is merely the reflective continuation in a hypothetical stance of action oriented to reaching an understanding (i.e. of communicative action), and argumentation is essential to the maintenance of communicative action (100), then the presuppositions of communicative action “are at least partly identical with the presuppositions of argumentation as such” (101). Hence the skeptic must retreat further, to pure strategic action.

However, Habermas claims, this is impossible. Such a retreat is just conceivable in the abstract, once a person has been formed, but the central role of communicative action in forming persons within society means that to shun communicative action from birth is to never become a person in the first place, and to withdraw once formed is to risk “schizophrenia or suicide” (102). Thus, although it is abstractly possible to avoid the presuppositions of argumentation (in the way that rocks do), it is impossible *for us*. There is no alternative.

In Habermas’ own summary:

My programmatic justification of discourse ethics requires all the following:

1. A definition of a universalization principle that functions as a rule of argument
2. The identification of pragmatic presuppositions of argumentation that are inescapable and have a normative content
3. The explicit statement of that normative content (e.g. in the form of discourse rules)
4. Proof that a relation of material implication holds between steps (3) and (1) in connection with the idea of the justification of norms.

Step (2) in the analysis, for which the search for performative contradictions provides a guide, relies upon a maieutic method that serves

- 2a. to make the sceptic who presents an objection aware of the presuppositions he knows intuitively.
- 2b. to cast this pretheoretical knowledge in an explicit form that will enable the sceptic to recognise his intuitions in this description,
- 2c. to corroborate, through counterexamples, the proponent’s assertion that there are no alternatives to the presuppositions he has made explicit.

Substeps (2b) and (2c) contain unmistakable hypothetical elements. The description we employ to pass from knowing how to knowing that is a hypothetical reconstruction that can provide only a more or less correct rendering of intuitions. (96-7)

We have seen that (U) fills the requirement of the first step. Alexy has provided the statement of the normative content in the form of discourse rules (required by step 3 and quoted above). These are an attempt to make the inexplicit explicit, following the process outlined in the expansion of step 2. As Habermas points out, this reconstruction may be in error, and empirical evidence can be brought to bear in order to test whether the reconstruction is adequate. Once an adequate statement of the rules is available, then (U) must be shown to be directly derivable from them. It is Habermas' claim that he has been able to show this. Whether these claims of Habermas are justified will be addressed in §6.24.

In this subsection, I have outlined the core of Habermas' discourse ethics. It is a fully-fledged meta-ethical theory, and it seeks to ground morality, not in a transcendental foundation (like Kant and Mill did), but in a transcendental-pragmatic argument that arises out of the practice in which we must engage in ethical matters. This argument has an essential core of empirical claims, through the "reconstructive science" of making explicit that which lies beneath unavoidable everyday human interaction. Ethical concerns (embedded in the lifeworld) are clearly distinguished from moral concerns (which have universalizable solutions). At a number of places in this outline, I have flagged points of concern. I will now turn to these, and look at some criticisms of discourse ethics.

6.2 A Critical Appraisal of Discourse Ethics

In this section, I will undertake two tasks. The first is a critique of the status of Discourse Ethics as a complete meta-ethical theory. Is discourse ethics able to carry the weight that Habermas places on it, as an explanation of the moral domain? There are a number of interrelated reasons for thinking that it cannot. Below, I shall investigate four of the weaknesses of discourse ethics: the assumption that discourse will inevitably reach consensus (including consideration of sharp cut that Habermas makes between issues of the right and the good; the problem of identifying normative correctness; the account of rationality that underlies discourse ethics; and the claim that discourse ethics is a purely formal and procedural meta-ethics, which does not rely on a substantive moral principle. Given that I can show that discourse ethics is flawed in these ways, then the second task, which I shall interweave with the first, is to explore what the insights of Discourse Ethics are that ought to be incorporated in an adequate meta-ethical account. While I shall identify a number of these, the synthesis of these

insights from discourse ethics with virtue ethics, under the modifying influence of my account of reasonableness, shall be deferred until Chapter Seven. This will provide the theory of moral development on which to base my discussion of moral education in Part III.

6.21 What is a Consensus?

As we have seen above, the idea of consensus plays a central role in Habermas' theory. In the ideal practical discourse, all will eventually give their approval to the norm supported by the better argument. In Habermas' characterization of consensus, all those affected come, through an exhaustive process of argumentation, to see that a certain line of argument is better than all the other lines advanced, and so all freely accept that its normative conclusion is the single best answer for each of them, considered individually. Consensus is a positive agreement by everyone that each norm is precisely what they would choose over all other possible norms. As he says, "it follows directly from (U) that anyone who takes part in argumentation of any sort is in principle able to reach the same judgments on the acceptability of norms of action" (121).

It is not clear why this must be true. All that follows directly from (U) is that we can conceive of a situation in which all agree to accept a particular norm. Underlying (U) is Habermas' assertion that, when we enter argumentation, we must be aiming to convince the others, using good reasons: "argumentative speech is a process of communication... [which has the] goal of reaching a rationally motivated agreement." This itself is not obviously true to everyone: it can be asserted that the goal of argumentation may merely be to reach mutual understanding, with consensus as a possible outcome. I shall not try to adjudicate on this claim, for even if it is granted, it is hard to see how the fact that all involved are *aiming* at a consensus guarantees that they will *reach* a consensus.

It seems to me that there are three possible outcomes to rational argumentation, even in the ideal: consensus, compromise² and a failure to reach agreement. Here is how Habermas (1996, 165-166, italics added) distinguishes the first two:

Compromises provide for an arrangement that (a) is more advantageous to all than no arrangement whatever, (b) excludes free riders who withdraw from cooperation, and (c) excludes exploited parties who contribute more to the cooperative effort than they gain from it.... Whereas a rationally motivated consensus rests on reasons that convince all parties *in the same way*, a compromise can be accepted by the parties each for its own different reasons.

² Habermas later (p 281) makes it clear that he thinks compromise "may well allow for strategic interactions." However, there seem to be no reasons for thinking such a compromise cannot be reached even in the ideal situation, with no coercion.

Consensus requires more than an unforced agreement: the reasons that lead to agreement must also convince all participants “in the same way.” Compromise occurs when some (maybe all) are willing to settle for less than they would ideally want on the grounds that, otherwise, no agreement could be reached. The exhaustive process of argumentation may lead all affected to a position where they come to see that there is not going to be a single version of the norm that each can accept in a strong consensus. Nevertheless, they may *each* see that, in the light of this stalemate, there is a norm which *all* are happy to accept as the best achievable outcome. It does not represent the ideal conclusion that each individually would prefer. However, it does represent a conclusion that each is happy to accept, given that irreconcilable difference has prevented each from convincing all others of their own preferred norm. Each accepts the norm as the best achievable, given that they must all live in community. Of course, it seems perfectly possible that, even with the best will in the world, an ideal discourse might come to no agreement on the contested norm at all, if different parties dig their heels in on “unnegotiable claims.” If I am right, and the outcome of an ideal practical discourse could be any of these three, then to aim at consensus does not entail reaching it, even in the ideal.

There are a couple of things to note about Habermas’ assertion of the inevitability of consensus (in the ideal). Firstly, it depends very much on the acceptance of a universalist model of reason, and I shall explore this connection in §6.23. Secondly, Habermas uses the word “norm” only to apply to statements regulating matters of justice (or the right). He says:

The universalization principle acts like a knife that makes razor-sharp cuts between evaluative statements and strictly normative ones, between the good and the just.
(104)

This, he says, is because only statements concerning justice can be fully universalized. Evaluative statements about the good life are so embedded in the culture and the forms of life of individuals that there is no possibility of reaching consensus on them. This is easy to see in endless and undecidable arguments about the good. Habermas requires this sharp cut because it is only thus that he can establish, through his transcendental-pragmatic argument, that a neo-Kantian account of morality is broadly accurate, in advance of entering into practical discourse.

Compromise is not capable of cutting the right from the good because it is not “razor-sharp” (see also Habermas’ discussion on pp 86-91).³ A compromise can be reached

³ It is worth noting that Habermas has attracted criticism for making over-sharp cuts in other areas as well. For example, Cohen (1995) claims that his distinction between the public and private spheres fails to appreciate the fluidity between them, while Meehan (1995) attacks his sharp distinction of ego-identity from moral identity.

in Habermasian ethical dialogues, where we recognize that we have conflicting ideas of the good life, and that it is acceptable for each of us to pursue our own.

If it is the case that, in some matters of the just and the right, a compromise is the best that can be achieved, then this amounts to an admission that practical discourse will tolerate a norm that allows some leeway to those with different conceptions of just action to pursue different ends. Such a case would arise if two proposed moral norms were incompatible, and some members of the discourse claimed A should override B, while others claimed B ought to override A. Such a pair might be competing normative accounts of fairness: “treat everybody equally” and “give everybody what they deserve.” An argument-driven agreement (that is, an outcome of a critical discourse) that the two could be reconciled through weakening each in relation to the other, in a way that leaves the supporters of neither side completely happy, but both sides content that this is the best achievable outcome, would be a compromise. If such a scenario is possible, then the sharp cut of right from the good cannot be made.

The distinction that Habermas draws between discourses of justification and discourses of application is designed to deal with this sort of difficulty. Habermas claims that strong consensus can be reached in discourses of justification, which deal with the attempt to redeem a contested norm. Such norms are highly abstract, generalized principles which, because of the abstraction from any concern about their application, can gain general acceptance (i.e. consensus). Thus we could agree in practical discourse to both of the principles: equality and desert. It is only when we try to apply them, to bring them back into the lifeworld, that clashes become apparent. But these discourses of application no longer seek the consensus of universal agreement:

The application of norms calls for argumentative clarification in its own right. In this case, the impartiality of judgement cannot again be secured through a principle of universalization; rather, in addressing questions of context-sensitive application, practical reason must be informed by a principle of appropriateness. (Habermas 1993, 13)

The problem with this, of course, is that the severe abstraction from situatedness in the lifeworld needed to achieve the abstraction of justificatory discourses strips meaning from the words that make up the proposed principle. In such a rarefied atmosphere, how do we decide what we mean by “everybody?” Who counts? What is the meaning of “equally?” In what respects need one be equal? These questions, and others like them, can only be decided by a consensus which is reached in close proximity to the lifeworld, and this requires consideration of the potential applications of the principles (see Warnke 1995, for a similar argument). Without such

situatedness, these purported moral principles become little more than “motherhood” statements.

Taylor (1991) is another who has asserted that, in order to formulate principles of justice, we need to have a conception of the good life. Habermas clearly recognizes this, as questions of justice in his scheme can only be decided by a community, each member of which is situated in the lifeworld, each with their own individual conception of the good life. He asserts, however, that a mutual agreement will be reached to drop questions of the good from the discourse, as they cannot be decided by consensus. Participants will agree, without coercion, to leave them to one side. Yet, as Taylor points out, it is far from clear why just these considerations will be dropped:

We have to do here with a central area of moral problems which focuses on weighing up the often mutually competing claims of the different virtues against one another and bringing about a uniform, consistent form of life. Questions of such a kind are at the centre of a moral life; it would therefore be completely arbitrary and unfounded simply to exclude this whole area by adhering to the false thesis that only virtues oriented towards charity have a moral character. (32)

There seems to be a circularity in Habermas’ argument. On the one hand, we have to assume that it is possible to divide all matters of moral concern into matters concerning justice or matters concerning the good life, if we are to be able to assert that only argumentation concerning justice will (inevitably) lead to a consensus in a practical discourse. On the other hand, the assertion that argumentation is solely about the inevitable force of the better argument, which guarantees coming to a consensus,⁴ is used to show that (U) can make the required cut of the right from the good.

Discourse ethics rests on a rational reconstruction of a number of presuppositions (§6.1). Habermas says “all rational reconstructions, like other types of knowledge, have only hypothetical status... rational reconstructions of presumably basic competencies [should be put] to the test, subjecting them to indirect verification by using them as inputs in empirical theories” (1990, 32). Thus, we could break the circularity in one of two ways: we could find evidence that we can presuppose the right/good cut; or we could find support for the presupposition that argumentation concerning justice must reach consensus.

Attention to how the distinction between the right and the good operates within our society does not provide empirical confirmation that they are cleanly separable. Benhabib (1992, 184-185) points out (with examples) that a sharp cut of this sort is incompatible with our intuitions about the domain of morality, and claims that

⁴ That is, rather than also entertaining the possibility of mutual understanding amongst difference (coming either to compromise, or to an understanding of difference based in a failure to agree).

Habermas has confused matters of the identification of the moral domain with considerations of the kinds of justificatory constraints that apply to moral judgements. So, in the case of the distinction of the right from the good, the empirical test seems to fail. Some analytic distinctions between right and good can be made (I develop this view in §7.24), but the cut cannot be made “razor-sharp.”

The cut between the right and the good is meant to be able to tell us which types of questions will, in fact, lead to consensus *before* we conduct the discourse. Yet Habermas has claimed, as we saw in §6.1, that moral theorists ought not to anticipate the consensus, but merely make contributions to it. How can we anticipate what will lead to consensus? Ideal practical discourse is counterfactual, so we cannot test a claim that a certain class of question (e.g. about truth, normative correctness or evaluative correctness) would necessarily lead to a strong consensus by entering such a discourse. Claims such as these need to be either established through principled *a priori* argument, or to be made as conditional claims, supported by empirical studies of actual discourses. The transcendental-pragmatic argument of Habermas takes the first route, and I have already argued that it fails. Certainly the history of social and political dialogue provides anecdotal evidence that questions of the right are more likely to lead to some sort of consensus than questions of the good, but continued dissension about key questions of justice hints that we have no empirical guarantee of strong (or even weak) consensus.⁵ Thus the second way to break the circularity also fails the empirical test.

The relationship between the actual norms that have been accepted (tentatively) by the cumulative practical discourse of humanity to this point in history, and the counterfactual consensual outcome of the ideal practical discourse (i.e. the valid norms) is a tricky one. It seems to be possible that we have been, and still are, on the wrong track, as perhaps societies that accepted (say) the subjugation of women have been throughout most of history. Next, I shall explore the extent to which discourse ethics gives us grounds for confidence that *our* norms are valid, or at least approximating validity.

⁵ It is interesting to note that several empirical studies (e.g. Frohlich and Oppenheimer 1992; Chan 1998) have been made to test the support for Rawls' (1973) claims about which principle of justice would be chosen by people in the original position - and that they have not borne out Rawls' predictions. Rawls makes substantive claims, of course, and Habermas would no doubt argue that these results reinforce the need for submitting claims to widespread discourse. Yet if his transcendental-pragmatic argument for the right/good cut fails, then his claim needs to be submitted to the discourse as well.

6.22 How Can We Anticipate Normative Correctness?

Habermas, we recall from §4.123, clearly distinguishes between practical discourses (concerning normative rightness) and theoretical discourses (concerning states of affairs). Nevertheless, he also refers to parallels between the two, and a little closer attention to some of the similarities and contrasts might help to clarify the idea that a practical discourse gives us grounds for confidence that our conclusions approach moral correctness. There are problems with this assertion.

For Habermas, that which underlies normative correctness is not objective: it is a product of the intersubjective world. Consider the way Habermas contrasts truth with rightness (normative correctness), on p 137: the function of truth is “the representation of states of affairs,” which are “in accordance with *the* world of existing states of affairs,” while the function of rightness is “the maintenance of an interpersonal relationship... in accordance with... *our* world of legitimately ordered interpersonal relationships.” These are different “*relations to the world*” (all italics in the original). On p 61, he says that “while there is an unequivocal relation between existing states of affairs and true propositions about them, the ‘existence’ or social currency of norms says nothing about whether the norms are valid.”⁶

Until agreed upon, it would seem, normative correctness does not exist, and since that agreement must be the same as that which would be the product of a (counterfactual) ideal practical discourse, we cannot be sure that it has yet been (or ever will be) created - especially if, as I argued in §6.21, we cannot be assured that the discourse would issue in consensus. Not only, in other words, is there no standpoint between language and reality, but social reality does not even exist until it is intersubjectively, and counterfactually, constituted. This leads to some problems that I shall address next.

Any actual discourse - institutionalized, overarching or whatever - must be finite. However, the consensus that constitutes normative correctness must be reached in an infinite discourse. The question thus arises: why ought we to put any weight on any outcome of any particular practical discourse? What norms are we to act on? Habermas’ answer is that the real discourses that have taken place over time and space give us an approximation of the valid norms on which we would, in the ideal practice discourse, come to a consensus. In Habermas’ words (previously quoted in §4.124):

⁶ This all strongly suggests that Habermas has a realist ontology for the physical world, and that what underwrites his consensus theory of truth for theoretical matters is the (unprovable) presupposition that theoretical consensus in some sense tracks the real world. I shall not explore this further here.

[In each discourse we] mutually assume an ideal speech situation. The ideal speech situation is characterized such that each consensus that can be achieved under its conditions is valid as a true consensus per se. The anticipation of the ideal speech situation is a guarantee that we may associate the claim to a true consensus with a consensus that has actually been attained. (Habermas 1971, 136; translated and cited by Horster 1992, 34)

We must just assume that our infinitesimally small fragment of discourse has tracked validity sufficiently so as to be reliable, because we anticipate an ideal practical discourse. Elsewhere, Habermas (1996, 14) says that “because no-one has direct access to uninterpreted conditions of validity, ‘validity’ must be understood in epistemic terms as ‘validity proven for us’.” We have to remember that when we act, for Habermas we act on the norms embedded in the lifeworld. These norms are “valid for us” because they have either never been problematized or, more likely, because they have been problematized in the past, subjected to practical discourse, and returned to the lifeworld. Are these norms at all close to being correct?⁷

In the end, according to Habermas, we must act on lifeworld norms, which have attracted wide agreement, as an act of faith: faith that the meta-practical discourse has been going on long enough, has been inclusive enough, and is free enough of distortions that we can tentatively accept its fallible outcome and act with confidence on it. Of course, if we have doubts, this tentative acceptance does not mean that we fail to continue to test these norms in practical discourse for their adequacy. It is this that makes for *critical* theory and practice. Habermas claims:

It is within the world that such learning realizes the conditions that must be presupposed as sufficiently satisfied for the unconditionality of the context-transcending validity claims. Here a certain degree of satisfaction counts as “sufficient” when it qualifies our current practice of argumentation as an exemplary local embodiment of the (unavoidably assumed) universal discourse of an unbounded community of interpretation. (15-6)... The universalistic meaning of the claimed validity exceeds all contexts, but only the local, binding act of acceptance enables validity claims to bear the burden of social integration for a context-bound everyday practice. (21)

We can judge what practices count as sufficient - i.e. which of the many unavoidable variations from the ideal can be ignored - by reference to the existential, teleological

⁷ There are some reasons for thinking so: claims that a good number of basic norms are widely accepted in all cultures: prohibitions on murder, stealing and dishonesty, for example (see Lewis 1946, 56-64, for a wideranging attempt); the advent of such agreements as the Universal Declaration of Human Rights, the spread of democratic institutions, and the increasing condemnation of slavery and many forms of discrimination. Yet there are also reasons for caution: we make these judgements from within a society that widely agrees with them, and our horizons (geographical, temporal and societal) might inappropriately bias our judgement; for every alleged advance, there are those who would claim the norms are flawed - the consensus is far from universal; other norms which are becoming more widely accepted are problematic (free abortion, bombing of civilians), and possibly not advances but backslidings; if we appeal to empirical evidence, the claim that the world is becoming a morally better place is highly contentious; the claim of an increasing moral consensus is based on a narrow, Western sample across a short time span, ignoring the lessons of history, present-day cross cultural comparisons, and subcultures within our own society.

and practical aspects of the ideal, as I discussed in §3.324. By presupposing the ideal, we can judge the actual, and take a stance on the reliance we are prepared to place on our existing discourses.

It is important to realize that any meta-ethical theory which does not entail some sort of direct and certain access to moral truth (through revelation or moral intuition, say) faces similar problems. If no such theory is tenable (and I argued in §5.1 that the multi-dimensional nature of morality gives us plenty of reasons for drawing this conclusion), then we need to be able to tolerate the consequent lack of guarantee to normative correctness. Virtue ethics, being multi-dimensional, has its own problems in specifying precisely what is morally correct. Yet this is not to say that moral inquiry is pointless; that we cannot ever make moral progress. All it says is that morality is not straightforward, not just a matter of following self-evident truths. Both discourse ethics and virtue ethics claim that through immersion in a culture, and through reflection concerning both ourselves and the moral views of that culture, we can gain moral improvement. Certainly, this is one of the key aims of moral education. In §7.23, I shall demonstrate that discourse ethics, through its dialogical emphasis, provides important resources to virtue ethics in this respect.

In Part III, I shall be considering the implications of this account for classroom discourses in moral education. In §9.1, I will consider how discourses about moral correctness assist the ethical development of children. In §9.413, we shall see that the interaction of lifeworld norms and moral discourses underpins an account of the avoidance of indoctrination. In §10.23, I will return to the question of moral correctness again, this time within the practical context of the community of ethical inquiry within the classroom. In each of these, Habermas provides the theoretical basis on which to build a robust account of the place of discourse in enabling and facilitating moral reflection: reflection which is especially essential in a pluralistic society within a rapidly globalizing culture.

6.23 Rationality

On the first page of the introduction to *The Theory of Communicative Action*, Habermas (1981) says: “The rationality of beliefs and actions is a theme usually dealt with in philosophy.... reason remains [philosophy’s] basic theme.” He readdresses the concept in an essay in *Postmetaphysical Thinking* (Habermas 1992) entitled “The unity of reason in the diversity of its voices.” As this title hints, the idea of reason/rationality is inextricably linked in Habermas’ thought with those of language and communication. Brand (1990, 138) says of Habermas’ work that “a lot hinges...

on the very concept of rationality.” In this section, I will claim that Habermas uses reason/rationality in three different senses, not always clearly distinguished.

My intent is not to give a full account of rationality, language and communication in Habermas, for such an undertaking would constitute a research project on its own. Rather, I wish to explore the place of rationality in discourse ethics, comparing and contrasting Habermas’ three concepts of rationality with my conception of reasonableness, as developed in §3.1. There are two main reasons for this. I have already questioned Habermas’ claim that ideal practical discourse will lead to a consensus in matters of the right. My analysis of the place of reason in this claim will support my rejection of discourse ethics as a sound meta-ethical theory. Secondly, if we have a clearer picture of the role of rationality in practical discourse, this will provide guidance to teachers in actual classroom dialogues for moral education in a way that enhances their students’ rationality (or reasonableness). In pursuit of these aims, I shall initially highlight a few distinctions that can help to clarify the senses in which Habermas claims that rationality is unitary and universal. This will lead to a survey of the place Habermas claims for the use of rationality in discourse.

6.231 Three Conceptions of Rationality

One of the difficulties in reconstructing exactly what Habermas takes rationality to be lies in his frequent use of alternative words and phrases for what seems to be the same concept. It is not always easy to work out whether he is using these as synonyms, or as means to make subtle distinctions. I have already noted the interchange of “rationality” and “reason,” but other words that seem to play the same role include “competence” and “competences,” “ability,” “reasoning,” and a variety of phrases in which the foregoing terms are combined with the adjective “communicative”: e.g. “communicative rationality,” “communicative competence” and so on (examples of each of these appear in the following discussion). Sometimes these latter phrases seem to be treated as subsets of a broader rationality, and other times they seem to be used interchangeably with rationality itself.

In order to assist in this discussion, I propose that we can distinguish between three separate but connected notions of rationality, as outlined below. I propose a terminology (not used by Habermas) to distinguish them. Each notion is used in Habermas’ writing in different places, and this is sometimes the source of confusion. I will substantiate these claims with detailed references below. Subsequently, I will explore whether Habermas may sometimes slip between these senses, leading to problems with his analysis.

Before investigating them, however, I need to point out another distinction that can be drawn. There is an ambiguity between rationality as some capacity (or cluster of capacities) that an individual has, and rationality as intersubjectively shared within a large group such as a culture (where it is likely that no single individual has a complete grasp of the whole). Habermas, with his sociological interests, is often concerned with the latter, which I shall label collective rationality; I, with my educational interests, with the former (individual rationality). While it is not always necessary to pay attention to this distinction, I shall at times have reason to raise it.

Presuppositional rationality: the anthropologically given fact that (virtually all, excluding perhaps those with specific neurological disorders) humans have a potential to engage in communication and argumentation, together with the presuppositions that this fact implies. Brand (1990, 18-20), referring to Habermas' use of the phrase "communicative competence" to mean what I have identified as presuppositional rationality, say that it involves:

the rules which are at the basis of linguistic communication and which competent speakers (that is, speakers with communicative competence) can apply although they would not be able to spell them out. The term communicative competence has been coined in analogy with the term 'linguistic competence' which was first used in a specific sense by... Noam Chomsky.... Thus communicative competence is the competence to achieve rationality with the means available to us in language, which, if used, are steering us in the way of reason. Those linguistic means do not only presuppose the possibility of discourse but also indicate what type of discourse we should engage in if communication is switched to that level.

Presuppositional rationality is thus a built-in ability to become rational through a reasonably normal upbringing.

Communicative competence is the ability to operate within these rules. This is an ability of the species as such. There is a unity in the plurality of language games, the unity pointed to in an answer to the question 'how a use of language oriented to reaching an understanding is possible'. Communicative competence is the capacity to handle language adequately as a certain tool in the necessary attempt to reach shared understanding. It is because this tool has certain characteristics that the attempt to reach shared understanding is not always doomed to failure. (ibid., 124)

Thus, all human beings (barring, presumably, those with certain neurophysiological malformations) are claimed to be communicatively competent in the same sense that Chomsky claims they are linguistically competent. That is, they are able, given the right environment, both to learn a language and to enter into communicative action, including discourse. In both cases, there is an underlying tacit knowledge (of universal grammar; of Alexy's rules of discourse) that everyone has, but which few can articulate, and even then only after considerable empirical study. I assert that it is presuppositional rationality that Habermas is referring to in passages such as the following:

I have gone into this controversy with the intention of rendering plausible a weak but not defeatistic concept of linguistically embodied reason.... Communicative reason is... [the] more or less trivial suppositions of commonality that make possible the cognitive, the regulative, and the expressive uses of language (Habermas 1992, 142).

For, although they may be interpreted in various ways and applied according to different criteria, concepts like truth, rationality, or justification play the *same* grammatical role in *every* linguistic community (Habermas 1996, 138).

Situated rationality: the actual capacity that any particular human has, at a given point in time, to engage in communication, and to think, rationally. This version of rationality has developed from the complex interplay of presuppositional rationality (as a potential and an in-built competence) with the social interactions of the child with others and the environment, influenced by maturational factors. As we shall see, Habermas' notion of situated rationality draws on the work of Piaget and Kohlberg, both of whom make universalist assumptions about cognitive development. I assert that it is situated rationality that Habermas is referring to in passages such as the following:

[Competences are] capacities to solve particular types of empirical-analytical or moral-practical problems (Habermas 1990, 33)

Ideal rationality: the ideal (and counterfactual) capacity that every person would have in the ideal practical discourse, or the ideal speech situation. This rationality plays the same role in each and every individual; i.e. it is that which leads from pooled opinions, desires, interests etc to knowledge of truth and normative correctness. Yet being an ideal, it is unachievable and "must not be... projected into the future as a utopia" (Habermas 1992, 145). Rather, it plays the same type of role as other ideals in Habermas' thought. I assert that it is the ideal rationality that alone leads to strong consensus in the ideal practical discourse, to which Habermas is referring in the repeated use of such phrases as "the unforced force of the better argument" and "rationally motivated assent."

Habermas has quite a lot to say about presuppositional rationality, and addresses ideal rationality at some length in a couple of places (as subsequent citations will demonstrate). In his scheme, both of these are universal and unitary. He argues at length for the universality and unity of presuppositional rationality, as we shall see below. On the other hand, he seems to assume rather than establish that all humans would counterfactually have a universal ideal rationality in the ideal discourse. I will argue that this is because he does not distinguish the two sharply enough, and because he inadequately analyses situated rationality, which must be the link between the two. Before expanding on this argument, I will further substantiate my claim that these three notions of rationality are to be found in Habermas' account.

Presuppositional rationality refers to the “general and encompassing structures of rationality” which “would have to be shown to be *universally valid* in a specific sense” (Habermas 1981, 137). Presuppositional rationality must be quite careful in its claims, for Habermas recognizes the weight of previous attacks on a substantive metaphysical concept of rationality. Hence, he seeks to defend a “concept of reason that is skeptical and postmetaphysical, yet not defeatist” (Habermas 1992, 116), later again described as “a weak but not defeatistic concept of linguistically embodied reason” (Habermas 1992, 142). In what sense, then, is Habermas’ presuppositional rationality skeptical and weak?

One of the results of these previous attacks that Habermas accepts is that we can no longer unproblematically assume that reason comes with substantive content. Thus he seeks to “explicitly state what we mean by procedural rationality after all substantial concepts of reason have been critically dissolved,” drawing a distinction between the contents of cultural traditions of reasoning and universal standards of rationality (Habermas 1981, 249). If we are to draw such a distinction, then what specifically remains under the heading of the universal standards of rationality? They are:

“the pre-theoretical knowledge and the intuitive command of rule systems that underlie the production and the evaluation of such symbolic expressions and achievements as correct inferences; good arguments; accurate descriptions, explanations, and predictions; grammatically correct sentences; successful speech acts; effective instrumental action; appropriate evaluations; authentic self-presentations; etc.” (Habermas 1990, 31)

It is important to note that presuppositional rationality *underlies* these productions and evaluations, rather than constituting them. I explored Habermas’ arguments for this claim in §6.1. To be actually capable of correct inference, good argument and so on requires a considerable degree of learning, and this learning must take place within a cultural tradition. Inference, argumentation and the others are the contents of reasoning, but it is important to note that they are also procedures. A sharp content-procedure distinction is impossible to maintain in reference to rationality, for as we acquire the tools of rationality, exploring the ways in which they can be used (i.e. treat them as content), we must also be simultaneously using some rational procedures to do so. The universal standards of rationality, then, cannot be the procedures of rationality, but that which universally underlies them. It is in this sense that Habermas’ presuppositional rationality is weak: it does not directly provide us with the tools of reasoning, but rather merely with that which underlies our capacity to acquire them. I suggest that this “pre-theoretical knowledge and intuitive command of rule systems” must be part of our biological heritage, as identified in Chapter Two.

As this acquisition proceeds, we become the bearers of a situated rationality. Indeed, the strength of Habermas' account of presuppositional rationality is that it provides good grounds for believing in an underlying commonality of all humans, preserving the possibility of communication between them, without necessarily privileging any particular historically situated version of rationality.

Situated rationality attracts relatively little attention in Habermas' work. It cannot be universal: each person will have developed their individual rationality to a different degree (on the unlikely assumption, which Habermas often seems to assume unproblematically, that it is a single dimensional capacity) or will have different aspects developed to different degrees (assuming that something like my account of multi-aspectual reasonableness is correct). That is, the detailed instantiations of individual rationality in different members of a single society will differ. Likewise, each society will develop a collective rationality that will differ from that of other cultures.⁸ Habermas clearly recognizes that individuals and societies do not exhibit anything like a perfect rationality:

It is, of course, obvious that the type of action oriented to reaching understanding, whose rational internal structure we have sketched above in very rough outline, is by no means everywhere and always encountered as the normal case in everyday practice. (Habermas 1981, 138)

As we have seen in §4.2, Habermas firmly bases his account of the development of rationality on that of Piaget and Kohlberg (Habermas 1990, 33ff). Built in to a Piagetian account are several universal and unitary assumptions: that development is a matter of moving from one stage of operations to another; that the order of stages is invariant; that all people everywhere go through the same stages; that the operations mastered in each stage are domain-general and context-independent (Donaldson 1978, 133-140). Each of these assumptions has been questioned, and many workers in child development would consider them to be, at best, oversimplifications, and probably false (see §3.23, also Sprod 1994c).

Yet, if these Piagetian assumptions are weakened or abandoned, we lose that thread of universality which constitutes the path for the situated development of rationality in individuals. As the title of the essay "The unity of reason in the diversity of its voices" indicates, Habermas is well aware that diversity is inevitable. The question is: "How

⁸ There is sufficient parallel between variations in rationality between individuals, and that between societies, not to be too concerned with this distinction. One reason why it might need to be taken more seriously would be if it were accepted that different cultures have incommensurable rationalities, for individuals in a single culture, given the constructivist account of rationality that I have argued elsewhere in this thesis, could not hold incommensurable rationalities. I believe, for similar reasons to those advanced by Habermas, that intercultural incommensurability in any strong sense is false, so I shall not explore the matter further.

much diversity?" A Piagetian account of cognitive development constrains the diversity to the results of differential development along a single path⁹.

There are, however, passages in this essay that support a more open interpretation of situated rationality, such as:

The procedural concept of communicative reason... discharges everything that has to do with content into the realm of the contingent and even allows one to think of reason itself as having contingently arisen. (Habermas 1992, 116)

Contingent factors will include differing cultural traditions of reasoning, different individual histories of learning, possibly even different preferred styles of reasoning or levels of native intelligence, which lead to a multitude of voices. On this account, presuppositional rationality, being purely formal, does not specify the content of situated rationality (and we need to recall that the content and specific procedures of reason are not easily distinguished). Indeed, the shape of the intersubjectively shared cultural version of reason, which has great impact on the situated rationality of its members, depends on its construction in communicative communities (as I argued in §4.22).

This is not to argue for radical incommensurability between different versions of rationality. Habermas argues, and I agree, that the presuppositions of communication are strong enough to keep all the variations of rationality within the bounds of mutual translatability (Habermas 1992, 116). Yet it is enough to establish that the development of rationality will not follow the same path in each individual's case. This would be true even if we restricted our concept of rationality quite narrowly to encompass only what I called the critical aspect of thinking in Chapter Three. It will hold even more strongly if we accept that rationality (or reasonableness, which I have argued is a better term) is construed as having many other aspects.¹⁰

With this analysis of situated rationality in hand, I will turn my attention to ideal rationality. If the progress through situated rationality could be seen as being along a single path, then there would be a single end towards which the path is heading. Even

⁹ I have simplified Piaget's claims somewhat here. In fact, he talks of sub-stages, makes allowances for some variation within developmental pathways and recognizes *decalage*, or the slipping between stages under some circumstances, amongst other complications. However, I do not believe that these simplifications affect the argument in any major way, and would only warrant reference to a tight network of near parallel paths.

¹⁰ I note that Habermas' account of rationality in *The Theory of Communicative Action* (1981, 8-22), and especially the explication of different types of discourses, has potential for development into a multi-aspectual account of reasonableness. However, as far as practical (and theoretical) reason goes, Habermas' account does seem to assume that rationality is constrained to what I have called the critical aspect of thinking. The other aspects, to some extent at least, seem to be consigned to other types of argument (e.g. the committed aspect of thinking is assigned to therapeutic critique; the creative to aesthetic criticism). I shall not explore this case any further in this thesis.

if the path were too long to traverse in a lifetime, there would still be an ideal rationality at the other end.

When he is drawing on Piaget and Kohlberg, there is a hint that Habermas believes the path is not too long. The following comment may be interpreted to mean that he sees formal-operational thought and postconventional moral judgment as the embodiments of ideal rationality in the theoretical and practical spheres respectively, and that such ideal rationality is within the reach of every young adult:

Piaget and Kohlberg describe the terminal competence of the young adult in the framework of rational reconstructions of formal-operational thought and postconventional moral judgment (Habermas 1990, 33).

However, it seems better to interpret these remarks as describing a well developed rationality, rather than a terminal one. It is not clear that Piaget would claim that formal-operational thought (or in Kohlberg's case, postconventional moral judgement) is "terminal" rationality, rather than a stage or style of thought that requires considerable specific learning of content (including procedural content) in order to operate better. For example, a formal-operational thinker for Piaget is capable of understanding formal logic, but does not, merely for that reason, know all of logic; that is, does not have the capabilities of a professional logician.

In *The Theory of Communicative Action* (1981, 24ff - see also my discussion in §3.323), Habermas gives a general account of the detailed abilities required by ideal rationality. The abilities required here are clearly far from the empty, formal potentialities of presuppositional rationality. There is a need to have firm situated abilities to adduce data, frame warrants, understand and motivate needs and values, anticipate consequences and so on. Again, these abilities are not beyond the reach of many individuals, but it is overwhelmingly probable that no actual individual is capable of using them all ideally. In any case, each of them is a high level capacity, dependent in turn on a multitude of lower level capacities and not constructed out of these in a unique way. These lower level capacities themselves can be more or less well mastered.

I have been building a case that rationality, for Habermas, is usually best interpreted as presuppositional rationality. This is the way the concept has been used in most of the references to rationality that I have considered. This is not to say that he does not address either situated or ideal rationality: at times, he certainly does. But at these times, it is noticeable that he is much less likely to refer to the universality or unity of rationality: these characteristics are by-and-large reserved for his "weak and minimal" presuppositional rationality. The question arises, however, as to whether, having

established the claim to universality and unity of presuppositional rationality, he also applies these characteristics to situated and/or ideal rationality. For example, he implies that an institutionalized public discourse (building a collective situated rationality) can approach or even achieve unified, universal ideal rationality.

The unity of practical reason can be realized in an unequivocal manner only within a network of public forms of communication and practices in which the conditions of rational collective will formation have taken on concrete institutional form (Habermas 1993, 17).

Implicit in this remark is a recognition that individual situated rationality is always somewhat fragmented and idiosyncratic. Practical reason cannot be unified at the level of the actual individual. Habermas argues that what is needed for the reunification of rationality in the real world of actual human beings is a public merging of individual rationalities. The unification of reason, he avers, can be achieved in a process of institutionalized collective rational action. It is important to note that there is some tension between this claim and the notion of counterfactual ideals. In the light of this, I will now turn my attention to the connections and pathways between presuppositional, situated and ideal rationalities.

How do these three categories fit with my analysis? Reasonableness, I claim, is equivalent to Habermas' individual situated rationality, though we have to recast the latter in terms of the five aspects. My analysis of how reasonableness forms supports the Habermasian alternative that "discharges everything that has to do with content [of reason] into the realm of the contingent," and that thus we all differ in the details of our reasonableness (though there is, given the constructivist account of the development of reasonableness, considerable overlap). Presuppositional rationality underlies our capacity to acquire reasonableness, and as such must be part of our biological heritage (Chapter Two). Ideal rationality is more problematic, and the next section will clarify this.

6.232 The Development and Endpoint of Rationality

One of my central concerns in this thesis is to give an account of the way teachers can, through the community of inquiry, help their students to improve their reasonableness. With this in mind, I will now consider the developmental pathway(s) between presuppositional, situated and ideal rationality. Some understanding of the development of reasonableness is essential to a teacher's pedagogical action in the classroom (§4.141). Can teachers assume that there is a single endpoint at which the development of reasonableness is aimed? If not, what can guide them? In the course of this investigation, I shall also consider the role such a convergence of reason plays in Habermas' thought. This will strengthen my judgement in §6.21 that consensus

cannot be taken as an inevitable outcome of ideal practical discourse, and hence that discourse ethics does not stand up as a complete meta-ethical theory.

In what follows, I will accept Habermas' account of presuppositional rationality. The use of language does imply that we must presuppose that communication is possible, that we are oriented towards reaching an understanding with others, and that this means that there is a necessary reciprocity to the process. It does seem to me that Habermas' argument that such a conception of rationality can secure the middle ground between a metaphysical myth and incommensurable radical relativism (Habermas 1992, 133-139) is sound. Presuppositional rationality also underwrites the teacher's assumption that all (normal) children are capable of developing reasonableness.

Now, there seem to be no problems in principle standing in the way of the development of an account of how individuals move from a universal, unitary presuppositional rationality to (a particular) situated rationality. After all, an identical potential can be developed in different ways depending on the environment in which it is developed. Consider again the parallel with Chomskian linguistic competence: all humans have it, but the differing circumstances in which it is developed lead to individuals being able to use different languages, being able to produce unique utterances and having differential abilities to communicate and understand. Further, Habermas' presuppositional rationality is only one part of the potential that a human has - a fairly minimal one. Other potentials will be present as well (not all of which need be universal - e.g. "native ability" or general intelligence), and the interaction between these potentials will differ amongst individuals. Yet it is this development which is of great interest to educationalists, and I shall return to it in §9.1.

Nor is there, on first glance, an obvious logical problem with the possibility of a move from differentiated situated rationality to a universal ideal rationality (though we must note that this is seen as a counterfactual move by Habermas). It is quite conceivable that different entities may start at the same place, move along different paths, and (potentially, dependent on contingent factors) end up at a common destination. This is effectively the position for which Habermas argues in asserting that a strong consensus is achievable in ideal discourse, and an ideal rationality would have to be created in this way. Indeed, given the intersubjective formation of reasonableness (§3.23 and §3.241), there will be some conformist pressures pushing in this direction: I shall look at these in §8.23.

Yet, although there is no logical difficulty with such convergence, there needs to be a solid argument as to why convergence will take place. After all, if different entities

head off in different directions, the safer assumption in the absence of such an argument would be that they are going to diverge to different destinations. Furthermore, there are a number of arguments that all lend credence to the conclusion that convergence is unlikely or impossible. I shall advance these below. Before that, however, I will explain why it is that Habermas needs to be able to establish that the various situated rationalities of concrete individual persons would converge on a single ideal rationality in a counterfactual ideal practical discourse.

6.2321 Why Situated Rationality Needs to Converge, in Principle

Although, as we have seen, it is presuppositional rationality on which Habermas lays most emphasis, the acceptance of a universal and unitary ideal rationality is vital to his argument that consensus is both possible and inevitable in ideal discourse in theoretical and practical matters. In order to see why this is so, let's have a closer look at the sort of consensus that is reached in the counterfactual ideal practical discourse. Here is Habermas on the idealizations of an ideal discourse:

the premise we posit for strict discourses, [is] that *in principle* a rationally motivated agreement must always be reachable, where the phrase "in principle" signifies the counterfactual reservation "if argumentation were conducted openly and continued long enough." (Habermas 1990, 105)

What argument leads us to the conclusion that agreement *must* always be reachable in principle? Habermas does not offer an explicit argument to this effect. A reconstructed outline of the implicit argument that underlies Habermas' conclusion seems to go like this:

Premise 1: Each participant in the ideal practical discourse has the *same information* at their disposal, because all possible participants (it was open enough) have contributed all that they want to it (it was long enough);

Premise 2: Each participant, at the time of decision, has at their disposal an *identical (i.e. ideal) rationality*;

Conclusion: Therefore, all will come to the *same conclusion*, because if identical data is processed by identical rational means, then there is one and only one conclusion that can be reached: the consensus.

Note that, although rationality has a clear role to play in this argument, just how this ideal rationality arises is left unclear. We saw above that Habermas asserts that we must "think of reason itself as having contingently arisen," and that this implies that real individuals each have their own different situated rationality. Yet if agreement

must be rationally motivated, by the end of the discourse this rationality must be the same in each.

Let's look a little further at the question of the nature of the participants in discourse. What do they bring to the discourse? In what ways are they the same, and in what ways do they differ? As we have seen, Habermas is critical of the position taken by Rawls, that each person, if placed in the "original position," is capable of making the correct decision. Being in the original position, Habermas argues, impoverishes the thinker. Though Habermas does not provide a single list of the ways in which the participants in a discourse differ from each other, we can try to reconstruct one from the comments he makes. In saying "for one thing, nothing better prevents others from perspectively distorting one's own interests than actual participation" (Habermas 1990, 67), he identifies that participants differ in their interests. A few lines later, he adds needs, wants and cultural values to the list. Presumably, they may also differ in emotional makeup, life histories, access to information and so on. In the ideal discourse, explications of each of these differences can be shared, and hence all have access to such information, as stated in Premise 1.

But on the question of rationality, he remains largely silent. Habermas, as we have seen, holds that all bring an identical presuppositional rationality to the discourse, for this is universal. It is, however, also weak, formal and empty. It is certainly not substantive enough to bear the weight of actually weighing evidence, engaging in arguments and reaching conclusions, because the abilities to do each of these, while they are *underpinned* by the presuppositions of rationality, are *actually learned* in social situations and dialogue. Consequently, the rationality that the participants bring to the ideal discourse must be either a situated or an ideal rationality.¹¹

At first glance, the former seems to be the more likely candidate. If this is so, and if the rationality that participants use to make the final decision is the variegated situated rationality that they brought into the practical discourse, then Premise 2 fails, and the argument does not hold. Differing rational processing of identical information will not necessarily come to identical conclusions. A modified version *might* hold, for it is not impossible that participants with differing, but overlapping¹² situated rationalities could nevertheless come to a consensus. There is no logical reason, as we have seen, to rule out differing paths to the same conclusion - but note that for Habermas consensus requires that the reasons convince "in the same way" (Habermas 1996,

¹¹ For a real discourse, of course, the rationality would have to be situated rationality. But then, real discourses do not have to issue in consensus.

¹² That the situated rationalities must overlap follows from two arguments: because rationality itself is formed in dialogue (the Vygotskian argument: see §3.23); and because it develops from a universal presuppositional rationality.

166, discussed in §6.1). I have argued that whether this is so is an empirical, and not an analytic, question, and needs to be decided separately in each case. Although we might find some cases where convergence occurred, this would still not be enough to establish the general case in the absence of a general argument. One does not seem to be available. And in any case, there are good reasons for believing that situated rationalities will not converge. Let's explore them.

6.2322 Will Situated Rationality Converge?

As hinted, an alternative argument is available to Habermas. This is that, just as (he claims) theoretical truth and normative rightness are the inevitable constructs of ideal discourse, then ideal rationality will also be the inevitable outcome of a discourse in which rationality itself is the subject matter. Thus, although all humans develop (and continue through life to develop, or at least alter) their situated rational capacity in idiosyncratic ways, in the ideal situation, all these different paths converge on a single Universal Reason.

On first blush, it might seem that, if we take a narrow conception of critical thinking as our model of rationality, then there is some plausibility to this position. This version of critical thinking sees rationality as little more than the application of formal logic to information, and logic might be seen as a self-evident axiomatic system. There are two reasons to doubt even this account of convergence on ideal rationality. Firstly, even if critical reasoning is implausibly collapsed to formal logic, there is no single agreed system of logic that can handle all logical problems; witness the development of deviant logics. Secondly, as I discuss in §3.121, Cherniak's (1986) work on minimal rationality concludes that "beyond generalizations about human psychology, there seems to be no transcendently 'right' kind of logician" (48), and that the idea of a perfect epistemic agent is "a profoundly inapplicable idealization, not just a harmless approximation of actual human rationality" (134).

This assumption of convergence becomes even more problematic if our model of rationality is multi-aspectual, as I argued that it is in §3.1, and based on "a rag-bag of 'quick and dirty' on-line stratagems" (Clark 1998, 179). Here are a number of reasons for thinking that different people's multi-aspectual situated rationality - that is, their reasonableness - is such that, even in an ideal discourse on rationality, convergence on ideal rationality is unlikely.

1. Talk of the unity of reason implies that rationality is a seamless, coherent whole. Yet the different aspects or heuristics of thinking do not essentially mesh with each

other. They may, taking the same information into account, push towards different conclusions.¹³

2. Two of the aspects of thinking are essentially located: the first in the individual history of the thinker (the contextual aspect); the second, in the person of the thinker (the embodied aspect).¹⁴ While it is true, *in general terms*, that these locations are the same for all persons, *in detail*, they are different. Thus the detailed instantiation of at least these aspects of reason in each person will be idiosyncratic, and will not “wash out” when we attempt to reconcile our different rationalities.

3. Our concepts, and even many reasoning processes, are themselves mostly metaphorical and thus not universal, as Johnson (1993, 32ff) has persuasively argued in relation to moral thinking. Because our key moral concepts and reasoning are too ambiguous and multivalent, the upshot is that the “view that there must be one and only one right thing to do in a given situation is, for the most part, mistaken” (187) - which is not to say that we cannot obtain “general guidance about how to live” (189).

4. Reasonableness is constructed through engagement in dialogue (§3.23). Drawing on Taylor’s (1995) notion of the dialogical self, Glaser (1998, 229) argues that, since the outer speech in which the emergent thinker engages is dialogical, then the inner speech of thought is also dialogical: “a pluralistic model of the individual thinker... gives full recognition to the fact that selves are not wholly consistent or unitary.” Since one of the major capabilities constructed through dialogue is reasonableness, then even within a single participant in a practical discourse, rationality is not unitary.

5. Warnke (1995) claims that reasons and justifications in discourses do not only aim at agreement: they are also involved in interpretation, which aims at insight and understanding. This applies to matters of the right as much as to matters of the good: “we might still have as many interpretations of the meaning of our norms of action and principles of justice as we have of our art and literature” (257). But while there are ways of establishing that some interpretations are better than others, often no one interpretation can be established as superior to all others. If this is the case, reasoning cannot be unitary.

¹³ Two examples. In the justice/care debate, cold, impartial logic (often associated with the critical aspect of thinking) will push us towards the former approach, while attention to our emotional attachments (under the committed aspect of thinking) will incline us to the latter (see Johnson 1993, 114-118, for a related discussion of Gilligan’s work in this regard). Secondly, within the critical aspect of thinking, we may strive to be both clear and concise, yet achieving the one may interfere with the achievement of the other.

¹⁴ Note that the emotions, from the committed aspect of reasonableness, depend on both the bodily reactions of those feeling them, and the contexts in which they have been learned.

For all these reasons, we must doubt the claim that individual situated rationalities will converge in what we might call rational discourse (i.e. discourse in which rationality is the content) towards an ideal Universal Reason. This has grave implications for the claim that discourse must converge on consensus.

6.233 Implications of the Plurality of Reasonableness

What is the conclusion we are to draw from this survey of rationality in Habermas' work? Firstly, let's return to the reconstruction of Habermas' argument for the inevitability of consensus in discourse (§6.2321). Through the assumption of convergence on universal ideal rationality, Premise 2 has reintroduced the monological subject into the ideal practical discourse. Habermas' monological subjects differ from Rawls' (§6.1), in that each represents their individual interests and retains their individuality in ways not available to the person in the original position. Habermas maintains that the pooling of information must be done interactively: "the revision of the values used to interpret needs and wants cannot be a matter for individuals to handle monologically" (1990, 68). Rawls reaches the information pool needed for decision by excluding these needs and wants from consideration. Once the pooling of information has taken place, however, they are the same as Rawls' subjects, in that they have become identical in their reasoning capacity. Now, any one of them could in principle come to the correct consensus decision.

Habermas needs situated rationality to converge on an ideal rationality so that consensus can be guaranteed in the ideal discourse. For all the reasons surveyed in §6.2322, especially those drawing on Cherniak, Clark, and Johnson's work, we have seen that it will not converge. Therefore, consensus cannot be guaranteed in any particular matter (though it is still possible, as are compromise and failure to agree). Since the strong distinction between the right and the good is underwritten by the guarantee of consensus for discourses concerned with the former (but not the latter), and since the guarantee of consensus requires a unitary reason (§6.2321), this is one more line of argument to show that the cut cannot be made "razor-sharp." Further, the insistence that all reasoners must, in the ideal rational discourse, become monological subjects with identical ideal rationality weakens Habermas' claim that discourse ethics is able to handle diversity without exclusion. Since the cut of the right from the good is central to Habermas' attempt to establish normative correctness through universalizability, and his dialogical proceduralism is designed to be inclusive of difference, these conclusions suggest that discourse ethics fails to achieve adequacy in its own terms. It is not satisfactory as a full meta-ethical theory.

Nevertheless, Habermas is right to identify reason as a central theme of philosophy, and a vital player in moral dialogue. Both from the Habermasian account of rationality itself, and from my reasonableness-inspired critique of it, arise a number of points that have great implications for classroom practice. Here, I will indicate what they are, and where they will be picked up in Part III.

In identifying presuppositional rationality, Habermas gives teachers a basis on which to build. The identification of the egalitarian, symmetrical and communicative nature of this underpinning mandates the use of communal approaches in the classroom, which value all, encourage participation and recognize difference (§8.21). This potential to engage in communicative action and discourse, present in every child, needs to become actualized into a burgeoning reasonableness. The Habermasian idea of a discourse that takes rationality itself as its content makes clear the central importance of inquiry (§8.22 and §10.22), and the merging of inquiry into a community (§8.23). Consensus, so central to Habermas, may not be the only goal of discourse, but it remains an important goal nevertheless. The concept provides important guidance to the classroom community of inquiry, and the teacher within it, as they seek to find out whether consensus is possible, when it is an appropriate aim and, if so, what the consensual position will be (§9.4 and §10.23).

The recognition that, contrary to Habermas, reasonableness does not converge on an ideal rationality highlights even more the need to allow and encourage students to represent their own interests in the community (§10.21 and §10.223). Nevertheless, reasonableness is a publicly tested and honed product (see also §3.241), and students need to expand their repertoire, and strengthen each of the aspects of reasonableness, through exposure to alternative approaches (§9.1).

Presuppositional rationality provides the basis from which Habermas' principle of universalization (U) is derived (§6.1) and it is time to address this last feature of Habermas' system.

6.24 Is (U) a Purely Formal Moral Principle?

Habermas claims that the basis of discourse ethics, (U), is a purely formal moral principle.

The justification of discourse ethics outlined here avoids confusion in the use of the term "moral principle." The only moral principle here is the universalization principle (U), which is conceived as a rule of argumentation and is part of the logic of practical discourses. (U) must be carefully distinguished from... substantive principles or basic norms, which can only be the subject matter of moral argumentation. (Habermas 1990, 93)

As we have seen in §6.1, (U) is derived from our rational reconstruction of the pragmatic presuppositions of argumentation. These presuppositions are made whenever we seek to convince another by using reasons, whether we realize it or not. They themselves have a normative content (Habermas 1990, 96) which must be made explicit in the form of the discourse rules, from which (U) must then be derivable.

So, (U) is merely a rule of argument, a principle of inference. It does not pre-decide what moral principles we will adopt, because they have to be proposed to, and accepted in, a practical discourse. Yet (U) itself has been derived from pragmatic presuppositions that *do* have normative content, and Habermas claims that these presuppositions arise solely in the attempt to convince another. Let's look again at the presuppositions:

(3.1) Every subject with the competence to speak and act is allowed to take part in the discourse.

(3.2) a. Everyone is allowed to question any assertion whatever.

b. Everyone is allowed to introduce any assertion whatever into the discourse.

c. Everyone is allowed to express his attitudes, desires and needs.

(3.3) No speaker may be prevented, by internal or external coercion, from exercising his rights as laid down in (3.1) and (3.2). (Habermas 1990, 89)

It is not hard to see that these presuppositions do have a substantive value written into them, that then transfers to (U). That substantive moral principle is respect for others.

For Habermas, this normative content of the presuppositions is built into language itself - a medium the purpose of which is to establish communication with others. But which others? Habermas has a circumlocution for the group of communicants but, as we shall see, it begs the question somewhat. The group includes "every subject with the competence to speak and act" (a few lines later: "all subjects without exception who have the capacity to take part in argumentation"). In other words, the others are those that are judged to have competence.

In §6.231, we saw that the best interpretation of Habermas' use of "competence" is to refer to presuppositional rationality. But it also brings us up against a second question: who decides who has presuppositional rationality? It is not uncommon to hear "those people are just not worth talking to" or "they never seem to understand a word I say." It is inescapable: the decision has to be made by us. Another way to put this is to say that the set of potential participants is those that we respect as potential communicants.

Benhabib (1992, 29-32) argues that we cannot derive a meta-ethical theory from a purely formal principle derived from the notion of performative contradiction. Instead, she claims, we must make two substantive ethical assumptions: universal moral respect (the assumption that all ought to participate) and egalitarian reciprocity (the assumption that all ought to have an equal right to speak, ask, be listened to and so on). Both imply respect. She recognizes that these assumptions do not tell us who is included in the “all,” and asserts that for some societies, the scope may be very restricted (e.g. my kin group), while for “modernity,” it extends to all humans.¹⁵

Taylor (1991) raises a related point:

The fact that I should argue with the aim of achieving domination-free understanding may admittedly be structurally implied by the logic of discourse. If, in other words, I attempt to assert my own interests irrespective of the objections other participants to the conversation raise, then I certainly violate the logic of the discourse. But why should I not do this? Why should I not attempt to reach my desired goal at the cost of being slightly inconsistent? (31)

For Habermas, to make the attempt to which Taylor refers in the last sentence is to catch oneself in a performative contradiction. For Taylor, it is merely a type of inconsistency. The participant in discourse needs a reason not to be inconsistent in this way, and what underlies the desire to not be inconsistent is that we see ourselves as “beings who should respect reason in themselves and others” (Taylor 1991, 30). Once more, respect emerges as a presupposed substantive moral principle.

Still the question remains as to who deserves this respect. The answer seems inevitably to be (as Benhabib saw) that we extend such respect only to those who are sufficiently like us; those with whom we feel we can genuinely communicate. Over history, this has left such varied groups as barbarians, children, women, underlings, servants and slaves out of the circle of respect - the Greeks left most of these groups out. Respect has been extended, not to all, but to the wise, the aged, the powerful, the clever, the word of a deity and so on. Universal respect has been a powerful idea in history, but it has not been the only account of who deserves respect.

Although such accounts seem somewhat foreign to the modern mind, it is important to see that there is no open-and-shut case for universal respect. Habermas says that all traditional ethical philosophies “have to protect a dogmatic core of fundamental convictions from all criticism” (Habermas 1990, 88), but the same might be said of universal respect in discourse ethics. Even then, there does not seem to be a sharp principled way of making the cut between those worthy of respect (as possessors of presuppositional rationality) and those not. Instead, the case for each group of those

¹⁵ In a lengthy and interesting footnote, she explores the extent of “all humanity,” and reminds us that communication extends beyond the use of language to gestures, cries and so on.

near the margins needs to be thrashed out in practical discourse, in which the subjects of the discussion may or may not be able to participate.

Since we need some substantive moral principle(s) to underpin discourse ethics, then it loses its purely formal character. According to its own requirement that all substantive principles need to be submitted to practical discourse and achieve a consensus before they are justified, then discourse ethics as a formal structure cannot be correct. But all this conclusion establishes is that discourse ethics cannot hold as a meta-ethical theory. It does not mean that we cannot reconstitute discourse ethics as a part of a wider theory.

Our ethics must start with the modernist lifeworld assumption of respect, and it must open this assumption to examination. This establishes a circularity, but it need not be a vicious circularity. As Benhabib (1992, 30) claims, we need to anchor ourselves in the lifeworld, but we open up the possibility of an historically self-conscious universality, checked against reflective equilibrium, and situated within our modernist historical horizons. Recall §6.22: we can never have a strong sense of certainty that any widely accepted consensus does actually track the norm which would reach consensus in an ideal practical discourse. Thus, I maintain, we are in fact in no weaker a position due to the collapse of a formalistic pragmatic-transcendental foundation. In both cases, we have to act on faith that our deliberations are giving us adequate, even strong, reasons for acting as we do. In both cases, we may be wildly mistaken. But in both cases, our very openness to argument and counter-argument, to dissent and agreement, to multiple points of view gives some confidence that there is every opportunity that sound conclusions will prevail and that unsound ones will be unmasked. This confidence, of course, cannot be shown to be (absolutely) justified, but we can argue that it is better than any of the alternatives. It seems to me that in both cases, *real* practical discourses rely on Peirce's cable, rather than a tight chain of argument:

Philosophy ought... to trust rather to the multitude and variety of its arguments than to the conclusiveness of any one. Its reasoning should not form a chain which is no stronger than its weakest link, but a cable whose fibres may be ever so slender, provided they are sufficiently numerous and intimately connected. (Peirce 1955, 229)

The arguments of this section have established that we must take respect to be a basic substantive norm if we are to have access to the insights of discourse ethics. We cannot start from a formal, morally neutral position. But equally, we cannot treat this foundational norm as an unproblematic foundation. Even while observing respect, we can investigate and possibly reconstitute it.

This theoretical account of the place of respect plays an important role in deciding the practical implementations of pedagogical action by teachers, both in the wider school (§8.1) and in classroom discussions (§9.2). It provides a justification for teachers to impose the values that arise out of respect even if the students have not had any opportunity to examine and adopt them. Equally, this account will be useful in my discussion of how teachers can avoid the charge of indoctrination (§9.41). These discussions, however, will appear in Part III. Before moving on to that task, it is time to look at the way in which the advantageous features of discourse ethics, as outlined in this chapter, can be incorporated into virtue ethics, discussed in Chapter Five, to form an appropriate target for moral education in the schools of the modern West.

6.3 Discourse Ethics: Benefits and Drawbacks

Chapter Six has undertaken three tasks: to outline discourse ethics, to examine it as to its adequacy as a complete meta-ethical system, and to prepare the ground for an analysis of the insights it can provide to a more satisfactory meta-ethics. We have seen that Habermas, working from an explicitly Kantian base, has claimed that discourse ethics provides a complete description of the nature of morality. However, in §6.2, I demonstrated that there are a number of problems with this account (the claimed inevitability of consensus, the split of the right from the good, the assumption that consensus tracks moral correctness, the account of rationality, and the claimed formalism). These, taken together, undermine Habermas' claim, and I shall expand on them further in §7.12.

Nevertheless, discourse ethics, when suitably modified, articulates well with virtue ethics: I shall defend this claim in greater detail in Chapter Seven. Discourse ethics requires the simultaneous and mutually supporting development of the intellectual virtues (in the form of reasonableness) and the foundational virtues of respect. Further, as I shall discuss in §7.25, in an ethics-oriented discourse, the virtues will be both the subject matter, and the products, of the conversation.

Discourse ethics' principle advantages lie in its dialogical and developmental emphases. In insisting that the development of human capacities and knowledge (both practical and theoretical) are dependent on interaction with others, discourse ethics leads us away from a faulty assumption that humans are somehow pre-existing subjects (see also the discussion of the philosophy of the subject in §4.141). Engagement with others becomes central to the development - even the possibility - of reasonableness, autonomy and moral virtue - three of the central themes of this work. Particularly when modified by the Vygotskian account of learning (§3.23), the

dialogical emphasis of Habermas' work takes development seriously. His account of immersion in the lifeworld also fits well with Aristotelian habituation.

Despite the critique of discourse ethics presented in §6.2, some of the Habermasian ideas investigated there do, with suitable modification, provide the resources for important conclusions. Although Habermas is (I maintain) wrong to claim that ideal discussions concerning matters of the right would reach consensus, nevertheless, the attempt to reach consensus, compromise or understanding is central to the construction of intersubjectivity and hence of reasonableness and autonomy. This will, as I explore in §7.23 and §7.24, contribute to my analysis of moral progress and justice. Further, and despite Habermas' failure to advance an adequate account of rationality, discourse ethics is of considerable use in clarifying the role that Aristotle allocates to practical wisdom and, even more importantly, it fills the gap in Aristotle's account concerning the origins of practical wisdom itself. Finally, a more careful analysis of the basis for dialogical action reveals that, despite Habermas' claims of formality, inculcated respect is one of the basic virtues upon which communication must be based.

I now turn, in Chapter Seven, to a more detailed look at the relationship between discourse ethics and virtue ethics.

Chapter Seven: A Meta-ethics for an Ethical Education Program

Chapters Five and Six have explored the requirements of a meta-ethical theory on which moral education can be based. In §5.1, I explored the connection between theories of moral development and meta-ethics. We saw that moral development is not the development of a single capacity (*contra* Kohlberg, who identified this capacity as that involved in making moral judgements on moral dilemmas). Rather, it involves the simultaneous development of at least five capacities.

Next, I engaged in detailed critiques of two meta-ethical theories: virtue ethics, through a survey of Aristotle's *Nicomachean Ethics* (§5.2), and discourse ethics, via a critique of the writings of Habermas (Chapter Six). In both cases, I identified the considerable strengths of the theories, but also some of their shortcomings. In this chapter, I will firstly recall, in summary form, the weaknesses and strengths of each theory (§7.1), and then suggest that, while we can use virtue ethics as our primary meta-ethical theory, discourse ethics is more useful in developing an account of the place of discussion in classroom moral education (§7.2).

7.1 Strengths and Weaknesses

In §5.24, I came to the conclusion that virtue ethics, because of its ability to handle multiple, potentially cross-cutting perspectives on morality, and because of its acceptance of the contextually situated nature of all moral matters, captures the varied phenomenological experience of our engagement in morality. As such, it can incorporate the insights of many other meta-ethical theories that reduce morality to a single principle. Like many meta-ethical accounts, virtue ethics recognizes the important facts that we are born into the world with a biological inheritance (§5.2), and that socialization plays a central role in building moral character (§5.21). I explored the intertwining roles of habituation to virtue and of practical reason in an adequate account of socialization, coming to the conclusion that, although Aristotle's account of virtue ethics emphasizes the importance of the development of moral persons, it does not provide an adequate account of such development (§5.22). In particular, there is a need for a fuller account of the place of dialogue in both habituation and the building of practical reasoning (§5.23).

Discourse ethics shows promise for completing this task. Yet as a competing meta-ethical theory, discourse ethics claims to tell the whole story of the moral domain, and its basic approach - a Kantian deontology - it at odds with the basic thrust of virtue ethics. Discourse ethics does assert a single core moral principle, derived from reasoning about the nature of language and communication, on which to base

morality, rather than appealing to a cluster of virtues. According to discourse ethics, morality is a matter of universalization rather than contextualized choice of the right option. Discourse ethics also narrowly defines morality so as to exclude questions that are situated in the lifeworld, whereas virtue ethics is essentially about living a good life (§6.1). Given these fundamental differences between the two, there can be no question of merely cobbling the two together. Instead, some of the basic assumptions of one or the other (maybe both) need to be questioned and reconstructed so that the elements of each that are to be combined can merge harmoniously.

I have argued in §5.24 that the basic premises of virtue ethics are sound, and in §6.2, I have indicated that the Kantian roots of discourse ethics are misguided. Consequently, it will be discourse ethics that will need to be modified most. As a preparation for this, I shall firstly (§7.1) gather together the strengths of discourse ethics - those features that we will want to preserve - and then (§7.12) survey the weaknesses that I have identified, which lead me to reject its strongly universalist and pragmatic-transcendentalist underpinnings. In §7.2, I shall show how this reconstructed discourse ethics can fill the holes in the Aristotelian theory.

7.11 What We Can Learn from Discourse Ethics

Although I have pointed above to the distinct differences between discourse ethics and virtue ethics, when we survey the strengths of the former, we find that there are some important intersections as well. Probably the foremost of these is the fact that both are social. They both recognize that the first steps in morality are taken in interaction with, and by means of becoming familiar with the practices of, other people. While Aristotle talks of habituation, Habermas refer to immersion in the lifeworld.

Discourse ethics places communicative interaction more centrally than virtue ethics, however, because it asserts that *any* advance in moral thinking must take place as a result of communication. Benhabib (1992, 53) suggests that “if there are certain moral and cognitive skills involved in reaching perspicacious, appropriate, sensitive and illuminating judgments... they bear a ‘family resemblance’ to the conversational skills and virtues involved in the ongoing practice of moral dialogue and discourse.” My analysis shows that this is no accident: it is precisely because Aristotle’s *phronesis* (i.e. the contextualized moral judgement referred to by Benhabib in the first half of her comment) is constructed in Vygotskian manner through engagement in moral dialogue (§3.23 and §6.232).

Discourse ethics’ emphasis on universalizability may not be strictly derivable from the presuppositions of communication in such a way as to provide a pragmatic-transcendental foundation (§6.24). Nevertheless, it does imply strongly that, behind

any attempt to communicate reasonably with another lies the presupposition that we must extend egalitarian respect to our conversational partners. Certainly, the idea of respect for others itself hinges on a concept of reasonable communication, but I have argued that this circularity is not vicious. In the absence of a foundation, our beliefs must be mutually supportive. Within the historical horizon of modernity in which we find ourselves, this presupposition of egalitarian respect extends to all humans, and if we were to ignore it, we would need to advance compelling reasons.

Discourse ethics thus provides a checklist for illegitimate practice - practice that denies respect to others. As virtues arise from within a culture, a virtue ethics account contains a danger of insularity and exclusion (§5.3). Reflection can provide a counter balance, but only if it extends beyond reflecting on whether one is properly instantiating the virtues oneself, to the matter of whether the versions of the virtues held within one's society are indeed virtuous. Such reflection is better accounted for by discourse ethics, and §7.23 will explore this issue further.

Habermas' clarification of the place of reason in ethics may suffer from a narrow construal of rationality (§6.23), but it does remind us that morality is connected with the giving of reasons, and the attempt to persuade others that our moral positions are correct. As Benhabib (1992, 50) points out, there is a danger of confusing moral cognitivism (that we try to justify with reasons our moral positions) with moral rationalism (that morality involves nothing more than a narrow rationality), yet to deny the latter does not refute the former. Moral development, as I pointed out in §5.1, is in part a burgeoning ability to handle moral matters cognitively.

Habermas also provides an important critique of the philosophy of the subject, throwing into question a number of assumptions that have shackled inquiry into the philosophical implications of the development of persons through their childhood (§6.1). Although his own discussion has been restricted to a Piagetian/Kohlbergian treatment of moral development (§6.232), and does not treat the growth of reason fully, thus being (in my view) inadequate, it opens the door to an inquiry which can be of the utmost philosophical importance. This critique demonstrates that we cannot take the individual as a starting point for philosophical investigation, ignoring the social constitution of that individual. Individuals cannot know their own interests in isolation, and what might initially be taken to be their own interests must be constructed, negotiated and clarified in conjunction with others, in community. Discourse ethics provides a way of understanding how such concepts as interests, needs, wants, principles, autonomy and solidarity are intertwined through their social interactive construction (Habermas 1990, 68). In Part III, I shall explore this in the context of the classroom.

Universalist theories have been commonly attacked for ignoring the fact that we are always already immersed in a situation. Habermas is well aware of this, and discusses the (as he characterizes it) Hegelian attack on Kant in great detail, notably in *Morality and ethical life: does Hegel's critique of Kant apply to discourse ethics?* (1990, 195-215). Consequently, he takes pains to emphasize that discourse ethics must be anchored in the lifeworld, that the moral questions submitted to practical discourse arise out of everyday practice through the problematization of the taken-for-granted. Practical discourse arises out of, and continues reflectively, ordinary communication (Habermas 1981, 25: see §4.124 and §6.1). Norms are not found in discourse, but in the lifeworld: discourse aims at their justification. Although the distinction between questions of the right and of the good is too strongly drawn in Habermas' work (§6.21), discourse ethics does provide a mechanism - practical discourse - for change and growth in moral understanding, and thus a way of escaping the potentially stifling conformity of immersion in the lifeworld (§7.23). Benhabib (1992) argues that discourse ethics can only approach universality through engagement with concrete others, by means of an interactive, and not a legislating, reason. She draws on Arendt's notion of "enlarged mentality" to show how extending our circle of respect results in a universalist perspective (§7.24). When practical discourse and enlarged mentality are combined with a more careful analysis of the emergence of reasonableness itself in communicative interaction (as I have developed in §6.23), this position has many implications for moral education (§10.23).

7.12 The Weaknesses of Discourse Ethics

In the discussion of discourse ethics in Chapter Six, I identified a number of its weaknesses. Here I shall summarize them.

Even if we accept the place of consensus as a necessary presupposition of argumentation, this does not establish that such a consensus is achievable in any particular subject matter, or substantive content. For confirmation that it is achievable, we need to turn to the empirical evidence. The empirical evidence of the long and multi-strand discourse conducted throughout history is that there are some questions (notably abstract questions of justice) where the hope of consensus seems to be a more viable target, and others (notably questions of the detailed individual instantiation of the good life) where it seems not to be. Nevertheless, we cannot assume that consensus is achievable on even the most abstract questions of justice, for such a universal consensus has never been achieved. Further, it is not clear that, even if we granted the hope of an ideal agreement in matters of justice, why that agreement needs to be a consensus rather than a compromise (§6.21).

Certainly, as such questions get less abstract and decontextualized, even the rough degree of consensus that adheres to some very general principles of justice gradually erodes. The possibility that agreements in matters of the right will be compromises rather than consensus supports the contention that the right cannot be sharply cut from the good, as Habermas maintains (§6.21). Indeed, I surveyed a number of other reasons for questioning the sharp cut of the right from the good, amongst them Habermas' narrow account of rationality (§6.23).

Even if we were to grant that ideal practical discourse would lead to consensus, the finitude of the discourses to which we have access implies that there are grounds for caution as to whether the norms which we presently accept track that consensus (§6.22). We need to note, however, that this is a version of a problem that will occur for any meta-ethical theory which does not propose some sort of direct access to objective moral truth. Nevertheless, it is not merely enough for us to be able to participate in discourse - it is also important to be able to engage in a meta-discourse which can evaluate the reliance that can be placed on our conclusions. Recognition of the fallibility of our currently accepted norms ought also to mean that we remain open to opposing views and (sufficiently) tolerant of difference. This, of course, raises the question of the limits of tolerance. I shall argue in §7.23 that discourse ethics provides us with the resources to address this issue, and return to the question in the context of the classroom in §9.2 and §10.23 .

My clarification of the three senses in which "rationality" (§6.23) is used by Habermas shows that the argument for consensus is based on a universality and unity of reason that cannot hold at the level of the actual situated individual. I am inclined to accept, with minor caveats, the case for the universality and unity of presuppositional rationality. Any attempt at communication does entail an attempt to convince, or at least to reach understanding, but it does not entail that consensus must follow. In the third sense of rationality, the argument that consensus would necessarily be reached in ideal discourse relies on a convergence of reason on a universal and unitary rationality. I believe, for reasons that I outlined in §6.23, that reason will not converge, and that the assumption that it could smuggles a monological individual back into the argument.

Finally, I have argued (in agreement with others, such as Taylor and Benhabib) that we need to accept that one or more substantive moral principles underlie discourse ethics (§6.24). Minimally, discourse ethics rests upon a principle of respect, and there is an unanswered question as to the scope of this respect. Since we need to take for

granted, at least initially, a substantive moral principle, discourse ethics cannot be a purely formal system. This is not to argue that the principle of respect is an unquestioned given: on the contrary, both it and its scope must themselves be submitted to practical discourse at any time they are questioned. Habermas claims that the key to the pragmatic-transcendental “foundation” for discourse ethics is “the fact that there are *no alternatives* to these rules of argumentation...; the rules themselves are not being *justified*.” (1990, 95, original italics). Yet these considerations show that an alternative, in which respect is circumscribed, is available.

From the considerations presented above, we must draw this conclusion: discourse ethics is not an adequate complete meta-ethical system. It relies on an equivocation between the presupposition of consensus as an aim, and the certainty of the achievement of consensus as an ideal outcome. It makes an untenable cut between the right and the good. It assumes a closer correspondence between current discourse and ideal discourse than is justified. It relies on an assumption that the multifarious situated rationalities of individuals would converge on a single ideal rationality. It does not establish a pragmatic-transcendental foundation, but rather needs to be anchored in the lifeworld by some substantive commitments. Consequently, if we are to maintain the many strengths of discourse ethics, we need to either supplement it, or incorporate it into another meta-ethical approach. I shall take the latter path.

7.2 A Meta-ethics for the Classroom

Now, I will sketch in a meta-ethical theory that draws upon the strengths of both discourse ethics and virtue ethics. Given the constraints on this thesis, I shall attempt neither to either fill in all the detail, nor to fully defend the theory. Nevertheless, my account will provide the basis for an extended exploration in Part III of the justification for, and the role of, discussion in classroom-based moral education.

As indicated, my meta-ethical approach will be based in virtue ethics, as providing a general account of what it is to be a moral person. A modified discourse ethics provides an account of the role of reflection and interaction within virtue ethics. Having investigated both theories in some detail earlier in this chapter, I do not propose to argue again for either the adequacy of virtue ethics as a base theory, or for the distinctive attractions of discourse ethics. Rather, I shall make four claims as to the advantages of incorporating discourse ethics insights into virtue ethics. These claims are that a virtue ethics theory that incorporates discourse ethics:

1. entails a need to inculcate certain basic virtues, especially the virtue of respect, without which the discursive part of a discourse ethics account cannot get going;

2. provides a much richer and more detailed account of the place of development, and especially of dialogue (communicative action and discourse), in ethics;
3. gives us an account of how virtue ethics can evolve, without being arbitrary;
4. enables us to incorporate justice as a virtue in a way that respects the universalizing claims of justice.

I shall now explore each of these claims in turn, before revisiting the role of virtue, both as the subject matter and as the product of discourse, in §7.25. The implications for classroom practice, particularly in terms of classroom discussions, shall be taken up in Part III.

7.21 Inculcation of Respect

We have seen (§6.24) that, although discourse ethics draws on the notion of a performative contradiction, in order to get a contradiction rather than just an inconsistency, communicative actors need to have a basic level of respect for those with whom they interact. As we saw, Benhabib identifies a number of substantive values that underlie discourse ethics, the major two of which are universal respect and egalitarian reciprocity. In what follows, I will use the term “respect” in an inclusive sense, to cover both of these, for we do respect others both in including them in the conversation and in ensuring equal opportunities to speak and be listened to. Indeed, in this more inclusive sense, respect covers other values, such as the positive attempt to encourage the more inarticulate to contribute (see also Benhabib 1992, 74).

However, I do not wish to try to map out the full extent of this notion of respect. Rather, I am pointing out that, however we unpack “respect,” it must include at least Benhabib’s two values. Some substantive moral principle of respect is necessary to underpin the ethical implications of discourse claimed by Habermas. Whether the particular instantiation of respect that we operate under is adequate can itself be made the subject of discourse.

In Habermas’ account, this respect is an inescapable feature of the use of argumentative language at all. I have argued, in effect, that this only holds for respectful argumentation, when the interlocutor is conceived of as being worthy of respect. In a society where some members are considered so inferior that argumentation is merely a means of imposing normative correctness, then respectful

argumentation is restricted to only certain members of that society. It is far from universal.¹⁶

The upshot of these considerations is that the possibility of seeing discourse as ethically important rests on an acceptance of universal respect, where “universal” minimally includes all humans who are capable of representing their own interests and giving and hearing reasons to some extent. That is, in the terms developed in Chapter Six, we need a sense of respect that extends to all humans who have developed at least some situated rationality.

Such an extension of respect is, I argue, a defining feature of modern Western culture. It forms part of the lifeworld within our culture. This is not to say that all members of our culture accept it: clearly, some do not. Nor is it to say that it can be unproblematically assumed to develop in members of our culture. Rather, it is an ideal of this particular culture, expressed, for example, in many of our basic political and legal documents. There is a presupposition of equality within the culture, and institutions or customs which seem to violate it must argue the case as to either why they do not, or advance compelling counter reasons that justify the violation.

I do not wish to enter the more contentious argument regarding whether some concept of universal respect either can be shown to, or ought to, hold in other societies. My intuition is that such a case can be mounted, though it will be a difficult one, seeing that it can be argued that imposing an ideal of universal respect on a society which does not fully recognize it, is to violate respect for the members of that culture (or at least, the members who already have respect within the culture). My interest in this thesis is not in the arguments for or against universal human rights, or for or against universal moral values. Rather, my interest is in moral education within our society - a society that has a general, even if somewhat vague and sometimes contradictory, lifeworld commitment to universal respect. We would expect that within such a moral education, questions such as the one about universality would be raised and contested.

In view of my concern with moral education within a Western modernist horizon, the practical question here is, at what stage and in what ways can we develop the virtue of respect for all others? Discourse ethics proper would give two answers to this question. Firstly, such respect is a taken-for-granted norm in the lifeworld of a modern society, although this is not to say that it always receives top billing, or that other taken-for-granted but contrary notions may not interfere with it. Secondly, its

¹⁶ Note that this is not a sneer from a perfect, modern society at more primitive societies. The same division exists in our society, and the boundary is contested. Does it include or exclude the brain-dead, fetuses, the senile, higher mammals? Some in our society claim that it excludes those with whom we strongly disagree.

force can be established in practical discourse, if and when its taken-for-granted nature is problematized. As respect itself is a complex notion, and there might be cases when either respect is not due (because someone's behaviour has forfeited a right to respect) or showing respect might entail treating someone in a way that may seem on the surface disrespectful (such as in Kant's argument that treating wrongdoers with respect entails punishing them: Rachels 1993, 135-136), such discourse will involve a clarification of the meaning and ramifications of respect. A virtue ethics account is quite similar: the virtue of respect must be habituated, though (as we have seen in §5.22) habituation itself requires reflection as it progresses.

The differences between the two accounts are small, as both mandate parents, teachers and other social actors to inculcate respect for others - by insisting on it, modelling it and other non-discursive means - and both mandate reflection on it. Thus there is no problem in incorporating a discourse ethics account into the virtue ethics theory. However, the differences, though small, are also vital. Virtue ethics provides a richer account of habituation to values (in Aristotelian language) through immersion in the lifeworld (to use Habermasian language) than discourse ethics does. On the other hand, Aristotle's account of reflection has little to say about how the moral subject becomes capable of such reflection, nor about how such reflection must be embedded within an intersubjective discourse. This is where discourse ethics' strength lies. Nevertheless, discourse ethics itself needs to be supplemented with a richer account of the development of reasonableness in becoming-persons, and this is provided by the Vygotskian account presented in §3.23. Let's consider this further.

7.22 Dialogue and Development

Virtue ethics, as I have outlined in §5.2 and hinted above, has an insufficiently articulated account of the role of reflection in the development of virtue. We have seen that Aristotle places a great deal of emphasis on habituation, and that his account of habituation includes reflection and the possibility that reflection is central to the changing of one's habits. However, Aristotle's account leaves mysterious how it is that reason itself develops: it is largely absent in childhood but present in adulthood.

Habermas' discourse ethics, as we have seen, avoids the worst of this characterization of reason through its denial of the philosophy of the subject. In failing to maintain a clear distinction between presuppositional, situated and ideal rationality, however, Habermas does place an over-reliance on the universality and unity of reason (§6.23). My account, which I have largely outlined in §3.23, pays close attention to the way in which reasonableness itself is intersubjectively constructed through dialogue. Thus it equips discourse ethics with a richer account of a simultaneous, interweaving and

mutually supporting development of moral insight and reasonableness. Both the moral insight and the reasonableness, being intersubjectively constructed, will have considerable overlap between children: in order to think well, children have to be able to think like others (§3.241). But both, however, being idiosyncratically constructed by each individual who will vary from the others in terms of history, family background, social situation, temperament, embodiment and other ways, will be in detail unique to the individual.

With the rectification of this weakness in place, discourse ethics provides us with a basis for explaining how children can bring forward for clarification, elaboration and possible modification both the moral virtues and the reasoning that they are assimilating within the lifeworld. Although in a homogeneous group, this may lead to the reinforcement of narrow views, the possibility of heterogeneity (and as we shall see in §10.23, this is partly the teacher's role) will often lead to the merging of wider horizons and the possibility of enlarged mentality - the subjects of the next two points.

7.23 Progress in Virtue Ethics

Virtue ethics was developed by Aristotle in a relatively closed and homogeneous society. In such a society, it is possible to believe that the virtues to be inculcated are uncontroversial and timeless. We, on the other hand, live in a society which can be described as multicultural, global, pluralistic and even (in postmodern terms) fragmented. Certainty about moral values is much less easy to assert unproblematically.

In such a society, one of the criticisms levelled at a virtue ethics account is that it mandates a return to narrow moral communities and traditional "certainties." In Benhabib's words, such communities can be "small, homogeneous, undifferentiated social units, particularly prone to intolerance, exclusivism, and maybe even forms of racism, sexism and xenophobia" (1992, 76). Of course, there is no reason why a community should not inculcate more inclusive and expansive virtues. The criticism is that, without a universal objective moral principle (or principles) to provide an ideal aim, a community will either merely reproduce itself (good virtues or bad), or drift into a different morality for no good reason.

Virtue ethics does, as we have seen, have some resources for refuting such a charge. Practical wisdom is responsive to context and needs to take circumstances into account. If the wider (societal) circumstances change, then practical wisdom "makes us take the right means" (Aristotle 1980, 1144a6) to pursue the good. The need for reflection so as to become the right sort of person, the sort of person who cares about their character, also contains within it the seeds of change (see §5.22 and §5.23).

Contextual sensitivity and personal responsibility for one's moral character make change possible. However, this virtue ethics account of change deals more comfortably with individual progress towards a moral character mandated by a particular society, rather than with societal progress towards a better morality (assuming this concept makes sense).

Two features of Aristotle's virtue ethics mitigate against an adequate basis for societal moral change and improvement. Firstly, the nature of the virtues themselves is taken somewhat for granted, so that reflection is more about how to hit the mark, rather than about what the mark ought to be. Secondly, reflection is seen to be a somewhat solitary pursuit, undertaken monologically by a solitary member of society, and concerned with that person's own moral improvement. Neo-Aristotelian work has grappled with these features (e.g., Baier 1994; Nussbaum 1994), but I shall turn to some of the insights of discourse ethics.

Aristotle's account of moral improvement can be usefully enriched by reference to the place of discourse in change, both individual and social. The previous two subsections have been concerned with the former, while this and the next subsection target social moral change. Of course, individual and social change are inextricably linked: an individual's new moral views may affect others and hence society as a whole, while being in a changing society changes many individuals. In §7.24, I will explore the place of justice and the right in virtue ethics, while here I consider the mechanisms of change in the accepted characterizations of the virtues.

In a pluralistic society, our taken-for-granted norms, inculcated in early upbringing and social interaction, will come into conflict with alternative formulations advanced by other members of the society. In this event, one of six things may happen: we may accept the other's norm; they may accept ours; we may both modify our norms to attain a consensus or compromise norm neither of us held before; we may both agree to tolerate the original plurality of norms; we may understand each other's norm but remain in non-tolerating conflict about the norm to adopt; or we may simply fail to understand each other and thus remain in conflict.

Given that the point of morality is to live in society with others without conflict (or at least to contain the harm arising from conflict), then each of the first four outcomes can be seen *prima facie* as local progress in morality provided that neither side's position was coerced. In discourse ethics terms, we have achieved an agreement in practical discourse (even if it is just the agreement to differ). I shall return to the notion of local progress soon, after considering further the implications of each of the six outcomes.

The difference between the fourth and fifth positions lies in the notion of tolerance. In the fourth outcome, we tolerate difference; in the fifth, at least one of us does not. Tolerance is a difficult concept in itself: it can be either a virtue or a vice, and the line between the two is not always clear. For example, few would disagree that to tolerate another's killing of a carrot growing in their own vegetable patch is a virtue, while equally few would dispute that to tolerate murder in the privacy of one's home would be a vice. Yet fierce arguments rage about whether to tolerate abortion is to tolerate the mere removal of a foetus or the murder of an unborn child, and hence whether such toleration is a virtue or a vice. As this example shows, even the language used to describe the act tends to prejudge the issue.

In claiming that each of the first four outcomes constitutes an agreement in discourse ethics terms, my account does not square with Habermas' characterization of the outcome of practical discourse. This is because the fourth outcome countenances the idea of an agreement on an irreducible plurality of norms, whereas Habermas would insist on the ultimate achievability of consensus on a single norm in the ideal.¹⁷ For him, the fourth position must, in the ideal, shift to one of the first three. This arises from his characterization of understanding:

We understand a speech act when we are acquainted with the kind of reasons that a speaker could cite in order to convince a hearer that he (the speaker) is entitled under the given circumstances to claim validity for his utterance.... Understanding an expression means knowing how one can make use of it in order to reach an understanding with somebody about something. Therefore... from the conditions for comprehending linguistic expressions [we see how] speech acts... are directed towards mutual understanding and thus towards a rationally motivated agreement about what is said. One would hardly know what it is to understand the meaning of an utterance if one did not know that an utterance *can and should* serve to bring about an agreement" (Habermas 1992, 78, italics added).

Again, we can see that Habermas has moved from the fact that in seeking to understand each other we have the presupposition that we might be able to convince, to the illicit conclusion that therefore we "can and should" (in the sense of *will*) come to agreement in the ideal. Thus, for Habermas, tolerance of difference is merely a way station on the route to (counterfactual, ideal) consensus.

When, in the absence of understanding, a speaker cites reasons for an utterance, the point is to motivate the other to change their view: to convince the other that the speaker is entitled to claim validity for the view, to paraphrase Habermas. So the initial view of the other has two aspects: the other does not already agree with the validity of the claim (otherwise there would be no need to convince them); and secondly, the other does not know the reasons why the speaker, "under the given

¹⁷ I'll leave aside here the distinction between a consensus and a compromise addressed in §6.21.

circumstances,” would want to assert that claim. Changing their view through the assertion of a reason, then, is somewhat ambiguous. If the other is to *understand* the speaker, then what changes is the other’s view that what seemed to be insupportable does in fact have some support, even if the other is unable to be motivated by that support to agree. If, on the other hand, the other is to *come to consensus with* the speaker, the other must change their view that the reason does not motivate them to accept the validity of the claim. These are two different views and, for the other, there is no reason why changing a view about whether a reason might motivate the speaker must lead to changing a view about whether the same reason would motivate oneself. While the attempt to gain understanding may entail that the speaker tries to convince the other, it does not entail that the attempted conviction needs to be about the validity *of the claim itself*, not that the ideal achievability of consensus must be assumed.

However we regard the status of the fourth outcome, it is clear that neither of the last two outcomes can in themselves constitute moral progress. The fifth, as it includes understanding each other’s position, may be a precursor to progress but cannot be moral progress in itself, if it does not decrease conflict. This is not to say that the position of one of the antagonists may not be morally praiseworthy, or even that each position may not have morally praiseworthy elements, but rather to say that no joint progress has been made.

Even local progress, of course, may not be seen more widely as moral progress if it involves agreement that others outside the local agreement would not accept. The fact that two hit men come to a local agreement that the right thing to do is to “rub out” an informer does not make this local progress into moral progress. We can make stronger claims that local progress is meaningful progress the larger and more representative the group in the local agreement is. This, I take it, is the key insight of discourse ethics. As we achieve wider agreement (either to adopt common norms, or to understand and tolerate divergent norms) with a wider circle of interlocutors, we can place more credence in the possibility that we are making moral progress. If we come across norms that we can understand but not tolerate, or even worse, that we cannot understand, then we must, if we wish to pursue moral progress, still attempt to reach agreement across this difference.

This process does not assume that there are universal objective moral norms. As we have seen in §6.22, even under the Habermasian account that postulates such norms “exist” as the counterfactual agreements of ideal practical discourse, we would not have access to them in any case. So moral progress can only be *prima facie* progress, through the achievement of wider local agreements, always subject to the proviso that

if we are exposed to a new, radically different point of view, we may need to rethink our norms and accept that our previous local agreement was probably mistaken.

Of course, just to become aware of difference is not enough to provoke change. On the contrary, it can evoke a siege mentality, where the awareness of difference leads to digging in at the present position. Discourse ethics requires a different response: a discursive attempt to redeem one or more of the normative claims expressed. This is to be done through the advancing and weighing of reasons. Even if no agreement is reached in the process of dialogue, and even if such a consensus is entirely beyond possibility, then it is still true that the discourse will almost certainly lead to better understanding. And although such understanding is not guaranteed to shift the initial positions closer together, nor to make the participants more tolerant of each other, it is more likely to do so than to do the opposite. The attempt shows respect for the other (§6.24).

Discourse ethics adds to virtue ethics a methodology for dealing with difference and disagreement; one which concentrates more on the *joint* attempt to construct moral values, and to create virtuous individuals, through striving for a virtuous community. Difference and disagreement can be resolved through the reaching of a consensus, under the influence of the better argument (if one exists), but they can also be resolved by reaching understanding and an agreement to tolerate the difference. In Habermasian discourse ethics, for matters of the right, the first is not only desirable but also ideally achievable, and its achievability marks out these matters as the subject matter of justice, the core principle of morality. Those matters in which such consensus is not ideally achievable are relegated to the realm of the ethical. If, as I have argued (§6.21), this sharp cut of the moral from the good is not sustainable, then we need a modified account of justice.

7.24 Justice as a Virtue

As I have argued in §6.24, given an underlying virtue of respect for others, to engage in practical discourse is to seek increased intersubjective agreement. If I respect you, and I offer you reasons in support of a contention, then I am hoping that the reasons will be compelling enough to convince you. As we saw in the previous subsection, this is somewhat ambiguous. I could be hoping to convince you either that my view is fully justified so that you ought to embrace it, or that my view is justified for me, given who I am. That is, I could hope that you and I reach consensus, or that we reach understanding.

Let's consider the first possibility: suppose for a moment that in offering reasons one must, at some deep level, desire full consensus. We have seen that even this cannot

demonstrate, as Habermas claims, that such an agreement is ideally inevitable (even if we restrict the domains of discourse to matters of the right), but this realization would not negate the desire for consensus. Of course, in giving my reasons I can realize that you might merely reply, “Well, I see what you mean, and why *you* might think so, but *I* don’t agree with you.” If I respond “Why?” you must offer your reasons, and in doing so, you hope to convince me. You may succeed, or I may respond to you as you responded to me. Here, we would have reached understanding but not agreement on the substantive contention. While we could agree that this is good, it would not have gone as far as either of us had (*ex hypothesis*) hoped when we entered discourse.

The question remains: Is the supposition with which the previous paragraph began justified? The supposition could have one of two natures: the claim could be conceptual, or it could be empirical. In other words, it could be that we must, as a consequence of the nature of reason-giving as an activity, have some (perhaps deeply buried) idea that consensus is the ideal outcome of that activity; or it could be that we are claiming that as a matter of fact, when people do offer reasons they do also (again, perhaps at a deeply buried level) hope to convert the other to their point of view.

The empirical claim would be difficult to prove. While it is couched in such a way that there is always a way out (we haven’t looked deeply enough), it just does seem to be the case that people do offer reasons for their views to others without seeking to convert them. This is commonplace in discussions of art, music and which football team to support, for example. These examples are all, it should be noted, from the ethical sphere, in Habermas’ sense. The first claim is also difficult to establish, for it itself seems to rest on further presuppositions (perhaps: that we must privilege our own well-thought-out views over those of others; or that unity is better than diversity) which can be coherently denied.

Consequently, I shall not try to defend either version of this claim here, although I feel that a weak version of the conceptual claim may be defensible. Instead, I shall examine the conditions under which people do, or do not, look for consensus rather than understanding. This will suffice for my later purposes. In doing so, I will consider the formulation I gave above for understanding: you understand me when you come to see that “my view is justified for me, given who I am.” Who I am, in this context, means how I am situated in the world as an embodied self, encumbered with likes, dislikes, interests, desires, a certain history and so on. Now clearly, you cannot know who I am in any full sense of the term. Such a knowledge is not even available to me, given that we are not transparent to ourselves. Yet, you must have sufficient access to who I am to be able to make the judgement that you understand (I note, without further analysis, that understanding cannot be full either).

Consequently, you need only to know about me in terms of those respects that are relevant to the validity claim in question. This gives a revised formulation for understanding: you understand me when you come to see that my view is justified for me, given who I am in those respects relevant to the view.

Let's apply this example to a claim about music. I claim that the Sex Pistols make the best music ever made. You dispute this claim. I offer reasons for the claims: I refer to a love of loud noise; to the elegance of rapid strumming and simple chord changes; to social conditions in the UK in the late 1970's and so on. In that you can imagine what I am like, and that you know (or come to learn) something of my history and so on, you can agree that you understand why I like the Sex Pistols, even though you are no closer to sharing that judgement. However, if you were more like me in the relevant respects (you also like loud noises, think rapid strumming and simple chord changes elegant and were anti-Thatcher), then you would be more likely to come to consensus with me, although you might opt for The Damned instead. If you were, *in the relevant respects*, just like me, we would come to agreement.

The upshot is that we have a greater hope of consensus when the participants in a discussion are more alike in the respects relevant to the claims under discussion. The point of Rawls's original position (§6.1) is that he claims *a priori* that matters of justice are such that the only relevant respects are those that any person would retain if they were placed in the original position (i.e. shorn of virtually all their personal attributes). Habermas contests this, and claims that any normative matter must be submitted to a practical discourse. In terms of this analysis, this is so that the participants themselves can decide what the relevant respects are, and then leave all the others out of the considerations. He does, however, seem to assume that in matters of normative correctness the upshot of such a discourse would be this: the only respects the participants in the discourse would count as relevant would be the action of rationality (identical in each) on shared information, thus ensuring a consensus outcome (see §6.232). In matters of the good, of course, other respects would be found relevant, and understanding, not consensus, would be the outcome.

If we do not follow Habermas' assumption, then the outcome of practical discourse must be found solely through real engagement in practical discourse. Yet we can predict that, on questions where the decision made in discourse is that many of the respects in which people differ are not relevant, and many of the relevant respects are shared by most or all of the discursive partners, then the hope of consensus can be stronger. Of course, we have seen that one of the respects that Habermas assumes is shared (i.e., rationality), is not, since situated rationality differs from person to person. Hence, consensus cannot be guaranteed. Nevertheless, the strength of hope

for a closer approach to consensus rests on a conception of persons, and how much commonality they have.

What I am seeking to establish is that there does exist a spectrum between, at the one end (the “goodish”), claims on which we can agree that understanding is the only appropriate hope, and at the other end (the “rightish”), claims on which there are reasonable grounds for seeking, and hoping to approach, consensus. In order to establish this, I need to achieve three tasks. Firstly, I need to show that there does exist sufficient common ground between persons (even persons who are inherently embodied and encumbered) so that, when we decide the relevant respects, we can have a reasonable hope that they retain sufficient similarity so that a meaningful consensus is a reasonable, even if not fully achievable, target. Secondly, I need to provide reasons for thinking that, while there are some claims for which we would be wrong to think that any sort of consensus is possible, there are others for which there are good grounds for thinking that an approach towards consensus is more likely. Thirdly, I need to show that there are good reasons for wanting to pursue wider agreement on claims of this sort.

Before embarking on the arguments for these three claims, I will reiterate that these arguments can be made with two target populations in mind. The more important one, clearly, is the totality of human beings, while the population to which I have restricted this discussion is the denizens of the modern West. My stronger claim is that these arguments may well work for the more inclusive population, but I shall be content if they hold for the more restricted population - the sort of society in which the schools I will discuss in Part III have to operate.

The first task I set myself will prepare the ground for the second. If I am to argue (as I will) that morality requires respect, that respect licences a separation along a continuum between rightish and goodish claims, in that the former can be discussed in a way that abstracts more (not wholly) from the particularity of people, and that this means that we can come to some discursive agreements that are not vacuous, then I need to show that even when many of the particularities of persons are laid aside, there still remain sufficient “relevant respects” in which the participants are alike for a substantive agreement to be possible. If the agreed stripping away of the varying particularities of participants left no common core, then it is hard to see how agreement could be reached on anything of importance.

In other words, the issue here is one of the content of human nature, or, more accurately, human universals. If a postmodern account of humanity, in which any individual is nothing more than a chance intersection of historical/cultural accidents

(Lyotard 1984), is correct, then no leverage can be gained for generalized agreement (or approach towards agreement) once all or most of these are bracketed. If, on the other hand, there is a sufficiently large and substantive core of traits common to all humans, such leverage is potentially available (although the actual achievement of consensus cannot be anticipated in advance of the discourse). An important root of the postmodern view lies in anthropological studies which assert the necessity to:

focus instead on the super-individual structures of language, ritual and kinship which make the individual what he or she is. Simply put, it is not the self that creates culture, but the culture that creates the self.... it also implies that nothing is "authentic," that there is no fundamental, originary nature of the human self against which we could judge a culture. (Cahoone 1996, 5)

Now, I do not wish to enter into a philosophical exploration of the self. Nevertheless, I would like to look more closely at just what anthropology does have to tell us about human nature and human universals. The distinction between these two I take to be this: human nature tells us what humans, being the biological beings they are, could not help but be like. Human universals tell us ways in which humans, as a matter of fact, are all alike. The former includes both genetically encoded traits and those that must inevitably arise through the interaction of a human person with the sort of environment a human must be in. The latter covers these, but also all those traits which, although contingent, are in fact shared through some process like diffusion into all cultures followed by cultural transmission. For my purposes, the latter concept is the relevant one, because it will inform us about the similarities that can survive the discursive stripping of particularities.

I shall draw on the work of anthropologist Brown (1991), who surveyed a great deal of anthropological literature in compiling his work *Human Universals*. There are three observations that I must make before looking at his findings. The first is that the focus of anthropology is on cultural groups, so Brown's lists of universals refer to those traits found in all cultural groups. Presumably this implies that most of them are found in all normal members of those groups, but it is quite possible that it excludes some individuals, such as the dysfunctional, maybe the very young or old, and so on. By and large, I shall ignore this complication here, for real discourses also have similar boundary problems in deciding who to include. The second is to note that because all humans exhibit some trait does not imply that all humans have the trait to the same degree, or developed in the same manner. As we have seen, to say that all humans reason does not imply that all humans have the same situated rationality. Thirdly, it needs to be noted that many of these commonalities are described at a high level of generality, and cover considerable diversity in the specific ways in which the commonalities are instantiated culturally in different groups. This last point will

complicate the analysis, and I shall consider its implications more fully after presenting the list.

In his Chapter 6, Brown devotes pages 130 to 140 to a fairly concentrated listing of human universals for which he has found considerable anthropological evidence, warning that the list is somewhat tentative. I have culled from this list a subset of universals (in roughly the order he gives them) that seem to me to have moral implications. According to Brown, the Universal People:

- interpret the external behaviour of others in terms of internal intentions;
- have the same facial expressions for happy, sad, anger, fear, surprise, disgust and contempt [thus presumably also having these emotions];
- use a smile to imply friendly intentions, and cry when they are unhappy or in pain;
- can mask or feign expressions;
- distinguish the self from others;
- can see themselves as both an object or a subject;
- see all persons as both subject to external actions and as partly responsible for their own actions, being able to distinguish between being in control of one's actions and not being in control;
- have the concept of intention;
- know that others feel emotions;
- can project themselves into other's states;
- live in groups, in which senior kin contribute to the socialization of the young;
- learn by copying and practice;
- cooperate in labor;
- undertake reciprocal exchanges;
- plan to maintain or manipulate social relations;

- have a form of government (in the sense that decisions which are binding on all are made);
- admire generosity;
- have laws (especially laws against rape, violence, murder);
- have a sense of rights and obligations;
- use punishment (including removal of the wrongdoer in some way) especially of acts that threaten the group;
- have more conflict in their society than they would like to have;
- understand seeking redress for wrongs, and consultation and mediation as means of trying to right wrongs;
- expect more cooperation from members of their in-groups;
- are not objective when comparing members of their in-groups to members of out-groups;
- distinguish right from wrong;
- have promises;
- empathize;
- have envy and ways to deal with it;
- have forms of etiquette and hospitality;
- have some standards of sexual modesty;
- have some food and word taboos;
- distinguish personal and group ownership of objects.

Bearing in mind the observations I made above, this is an impressive and far from empty list. As my third observation noted, it is also a list expressed in fairly general terms. Even if we know that all humans have a sense of rights and obligations, for example, this does not tell us what those rights and obligations are. Different groups can assert quite different clusters of rights and obligations. So a sceptic can claim that

even such universality as is picked out by this impressive list is too indeterminate to generate much in the way of hope for agreement in rightish discussions.

There seem to be two sorts of replies to this point. The first is to note that there must be sufficient commonality in what different cultures take to be (say) rights so that we can recognize that these are in some sense the same things as what we call rights. But this is unlikely to get us too far. Given a family resemblance account of concepts, the commonality may not have any universal core, and hence it may not be enough to generate optimism that some questions may tend more strongly towards agreement. Hence I will turn to the second sort of reply.

Among the other universals that Brown lists are poetic and rhetorical speech forms, metaphor, narrative and storytelling. Johnson (1987; 1993) has discussed the role that such imaginative structures play in human affairs, especially in morality. He, too, draws on the family resemblance account of concepts to assert that all our key moral (and other) concepts are prototypical: they have a core, relatively uncontentious in a given culture, and shade off radially into more and more contentious exemplars of the concept. These radial meanings are linked to the central meaning metaphorically.

So far, given the reference to the “given culture,” this account seems unhelpful to claims for some sort of universality. But Johnson’s *The body in the mind* is an extended defence of the claim that the most basic and central of these prototypes and metaphors are derived from the experience of the human body as it is situated in the human environment. These experiences are universal, and they underlie far more complex and abstract concepts. For example, the highly abstract moral notion of justice is underpinned by the ubiquitous bodily experience of balance (Johnson 1993, 194). Although our moral concepts are elaborated in a variety of culturally influenced ways (so that balance does not define justice in any straightforward way), we have now found good grounds for asserting that there is some commonality upon which to pin a reasonable hope of progress towards agreement in matters which tend towards the rightish end of the continuum. This agreement may (certainly not “will”) approach consensus, and is indeed more likely to than in matters towards the goodish end.

As Johnson (1993, 194) comments, “What is universal is likely to be so abstract that, taken by itself, it places only the most general constraints on action.” Yet it does provide, as he puts it later (p 237) “the common basis for a possible dialogue” across all humans. Because we must talk in general terms before we can plug into the commonality of humanity, we can understand why it is that more abstract and general claims hold out more hope for agreement. These claims are not vacuous, because they tap into a vast, even if general, reservoir of commonality. Nevertheless, deciding the

detailed application of generally held moral maxims will still be fraught with difficulties, because of the level of abstraction needed to approach agreement. I will defer this problem for the time being and turn to the problem of how we can distinguish claims for which consensus is a reasonable aim from those for which it is less reasonable.

In approaching my second task, I will draw on Brown's list to consider how morality is related to items on the list. Humans live in groups, and this implies that there needs to be some mechanism for mediating their interactions. Indeed, all humans plan to maintain social relations, distinguishing right from wrong, have laws to require right action and use punishment to enforce them. Indeed, Brown is more specific: all humans consider rape, violence and murder to be wrong. Although much more could be done in terms of relating morality to Brown's list, I feel that this is enough to establish that the institution of morality is ubiquitous in humans, and that it has a broadly determinate character.

At the start of this subsection, I referred to respect as an underlying virtue of discourse ethics. Does Brown's list give us any grounds for asserting that respect for others is a universal human virtue? We have seen that respect (in a broad sense) is a core moral principle of the modern West, even if it is not always observed. Certainly, schools as an institution have been based, since the introduction of the ideal of universal education, on respect for all citizens, and this will suffice for my purposes in Part III. Nevertheless, I will briefly consider whether Brown's list indicates that respect is a human universal.

To respect others is to allow that they are persons in their own right, with the sorts of attributes that we value in ourselves, taking account (insofar as we are able) of the ways in which they differ from us. All human cultures ascribe intentions to others and interpret their behaviour in those terms. They feel, and recognize, key emotions, including hurt and anger, and empathize, projecting themselves into the other's state. They ascribe responsibility to themselves and others in appropriate circumstances. They understand reciprocity and admire generosity. These characteristics begin to map out a broad concept of respect for others - but which others? The fact that all humans expect more cooperation for members of their in-group, and that they evaluate in-group members more highly than out-group members implies that, although respect for others is a human universal, the scope of that respect can vary quite markedly.

Let's turn to the question of whether accepting an underlying virtue of respect for others gives us any grounds for distancing rightish claims from goodish ones. To explore this, I shall narrow my focus from respect as a whole to a single aspect of

respect. Respect implies that we recognise that others, like us, do **not** like being harmed and do like pleasure (however their detailed understanding of harm and pleasure differs from ours). Now, to knowingly cause certain types of harm depends on knowing the detailed attributes of the person harmed. If I wish to harm you by playing Barry Manilow tapes at full volume, then it had better be the case that you have some musical taste, for if you actually get pleasure from Barry Manilow, the ploy will not work. On the other hand, if I wish to harm you by chopping your hand off, your detailed attributes seem to have little (but not nothing) to do with it. Almost everyone will be harmed by this act.

Some harms seem to be more basic than others, and the reason is that they are so general that they deprive the person harmed of the opportunities to seek other, more particular pleasures and avoid other, more particular, harms. Without a hand, you are deprived opportunities to do many things; with Barry Manilow playing, there are still many other pleasures open to you - and, mercifully, the tape will soon end. Because they are more general in this way, such harms are more likely to be agreed to be harmful by many differently situated people. Therefore, we may start to seek agreement on matters such as these with a greater hope of approaching consensus. These matters sit at the rightish end of the spectrum. Note, I am **not** asserting that this exhausts the criteria for distinguishing rightish claims from goodish claims; merely that at least one criterion for doing so exists. This is all my second task requires. Note also that I have based it on a notion of broad respect for all humans, so that the claim is limited to societies where such a principle holds.

So far my analysis has been restricted to an example where one human interacts with another. There is another dimension to interhuman interaction: how we are to treat others in general. This is where my third task arises. If it is established that I can cause you harm both by playing Barry Manilow tapes to you and by cutting off your hand, then respect for you (and morality) require that I do neither. The difference is, as we have seen, that I need to know some fairly particular facts about you in order to establish the former, whereas pretty common facts about humans in general will suffice to establish the latter. Having moral rules-of-thumb that tell us how to treat others can save us some work. Compare the following two possible rules:

1. All humans who dislike Barry Manilow ought not to have Barry Manilow tapes played in their vicinity.
2. All humans who dislike having their hand cut off ought **not** to have their hand cut off.

If I am to apply the first rule to you, I need to do some investigation about you, but it seems unlikely that anyone would ever bother to do such an investigation before deciding to apply the second. It could plausibly be replaced with:

2. All humans ought not to have their hand cut off.

Given that we must engage in many actions with moral import, having rules-of-thumb that cut down on the number of investigations we have to do before acting makes a lot of sense. These rules-of-thumb are more useful if they can be applied with less particular investigation. We have seen that claims which require less particular knowledge of the individual(s) to which they might apply are those that lie towards the rightish end of the continuum. They can, all things being equal, be applied to large numbers of people in the absence of particular knowledge.

Nothing I have said implies that rightish rules-of-thumb need always to be applied universally. Compared to goodish rules, they tend towards that, but they may clash with other rightish rules, or there may be unusual circumstances that mitigate against the rule in this particular case, or it might be unclear as to whether the situation is one that calls for this particular rule, and any of these will require further deliberation before the rule is followed. So situated decision is always needed in all moral contexts (even if it is routinized and largely tacit), and widely agreed, contextually impoverished rightish rules can ease the burden and assist in the routinization of the action. We can act from a sense of justice: that is, exhibit the virtue of being just.

To recap: Habermas claims, as we saw in §6.232, that there is a class of contentions that would, in principle, lead to consensus in practical discourse. This class maps out the domain of justice. I have argued that we cannot decide the members of this class, if it exists, in advance of actual practical discourse. It is an empirical and not an *a priori* matter. But this does not establish that there are no contentions that could achieve consensus, and I have argued that it is not the case that all contentions are equally unlikely to reach consensus. I have advanced grounds for asserting - rather than Habermas' sharp cut between contentions that can or cannot reach consensus, or the pessimistic conclusion that all statements of value are equally undecidable - that contentions range across a continuum of varying degrees of likeliness to attract widespread agreement.

In what follows, I will extend this argument. Firstly, I will illustrate this continuum by looking at some candidate pairs of contentions that differ in their position - one closer to the "wide agreement" or rightish end and another closer to the "personal value" or goodish end. Then I will look more closely at two difficulties (alluded to

above) with the assertion that the position of a contention on the spectrum carries any special status: the problem of acting on rightish agreements when agreement, though wide, is contested; and the problem of applying very abstract contentions in particular circumstances. I shall call on Johnson's work again to show how my modified discourse ethics can give a convincing account of justice based on widely agreed norms. Justice in this sense will be an important virtue, but not one that trumps all other considerations.

I have argued that some contentions that are submitted to practical discourse show greater promise of achieving wide agreement than others. For example, we seem more likely to reach an agreement that "Innocent people ought not to be harmed unnecessarily" than we are than that "Children ought to play Aussie Rules rather than rugby." We also seem more likely to reach agreement on abstract, highly decontextualized statements such as "All children ought to have an education" than we are to similar, but more concrete and contextualized statements such as "All 15 year old children ought to study *The Merchant of Venice*." In my terms, the former sit towards the rightish, and the latter towards the goodish end of the spectrum. A Habermasian would consider the former to be a matter of justice and the latter a matter of a person's preference for a good life.

I want to reiterate that there does seem to be a genuine spectrum here. If we compare the two statements at the personal value end, we can see that the football code contention would most likely gather even less agreement than the Shakespeare one. The evidence for this statement is that schools will countenance a syllabus that includes compulsory study of *The Merchant of Venice* where they generally do not impose footy and exclude rugby. Further, the degree of agreement is contextually sensitive. A community on the NSW-Victorian border would presumably reach less agreement on the football code contention than one in Melbourne, while one in Llanelli, Wales would most likely reach a high degree of agreement that it was wrong. Similarly, while the innocent persons contention has been likely to garner widespread agreement across time and societies, the education for all one will only have approached this in quite recent times.

From our historical vantage point we can see that some contentions have gathered wide (not universal) agreement both within the present global society and the sum of human societies over time. These more inclusive viewpoints have often been built in the attempt to incorporate the views of all, where possible. There are however, as I have argued (§6.21), a number of problems with extrapolating such agreement to an ideal consensus. I will now consider certain difficulties with the assertion that the fact that a contention has achieved a higher degree of agreement carries any special status.

The first of these is that, despite widespread agreement that certain contentions (let's use "all children ought to have an education" as an example) are just conclusions and so basic human rights, there remain those who are outside this agreement. While some of these dissenters are outside the West, others are members of our pluralistic society. This can be seen as a problem for discourse ethics, for if universal rights are only known to be established when universal consensus is reached, we would have difficulty claiming certainty about any human rights at all. Given the basic assumption of universal respect that underlies discourse ethics (§6.24), it might seem that its basic moral stance entails that we cannot impose our own conclusions about what does constitute a human right on others: that we must first seek their agreement. Yet such a conclusion would often lead to moral paralysis.

Of course, opposition to a claim of right in a real practical discourse does not necessarily denote a genuine belief that the claim is wrong: people can be duplicitous and self-serving. A dictatorship, or even a complacent majority, may suppress education in order to quell dissent. In an ideal discourse, of course, all must speak sincerely (see §4.123), but we have no such guarantee in any real discourse. And it is precisely those who duplicitously deny human rights to others for their own (immoral) ends who ought to be resisted and opposed in the name of morality. Yet, even though such cases exist, it is problematic to assert that all those who disagree with us must be duplicitous. At least three other possibilities exist. The others may be sincere but mistaken. It may be that neither side in this dispute is wrong, because we have a genuine case of competing values, neither of which has any hope of reaching consensus, so that the best we can hope for is mutual understanding and toleration. And of course it might be we that are wrong.

This discussion is leading into difficult areas of political and cross-cultural moral theory, and, given the constraints of my project, I do not wish to follow it too far. So I will round it off by stating, without detailed argument, a few tentative conclusions. It seems to me that moral paralysis is not a good thing. We must act on moral convictions, for not to do so risks allowing immorality to thrive. Nevertheless, the lack of actual consensus on all but the most abstract (and hence potentially empty) moral principles must also give us pause for thought. There also exists the danger of moral zealotry - seeking to impose on others a moral conviction which is not justified. This danger is also complicated by power inequalities. Discourse ethics tells us that the most important guide that we have between these twin dangers is the position of our convictions on the spectrum of agreement. Those moral contentions that achieve most agreement, particularly agreement across other differences, stand most chance of being justified. I cannot advance this as a sole criterion, for such a position faces

other objections. I will mention one: it is difficult to account for the lone voices that oppose what they see as widespread immorality (such as slavery) and eventually succeed. But here again discourse ethics offers assistance. It tells us that we must be tolerant, willing to keep an open mind and to treat our conclusions as potentially fallible; recognize that they bear a “time and knowledge index,” as Habermas puts it (1993, 37). It encourages listening and reasoned engagement across difference.

I’ll turn now to a second problem: one to which I alluded above in referring to abstract, possibly empty, moral principles. Allegedly universal moral principles are highly abstract and general. For Habermas, following Kant, this is because the ultimate moral principle is purely formal. This difficulty is addressed by Habermas’ distinction between justification and application (1990, 181-182; 1993, 35-39). Justification, he claims, is the process of assessing the proposed norm independently of its application in any particular situation, whereas a discourse of application assesses, not the norm, but the appropriateness of the norm to the detailed situation.

This clear distinction between justification and application seems to be another of Habermas’ unsustainable sharp cuts, closely related to that cut between the right and the good. To even make sense of the proposition “All children ought to have an education” requires interpretation of the key terms, including “children” and “education.” Such clarification cannot take place in isolation from the particulars of the lifeworld and its understandings. As Benhabib (1992, 23-26) makes clear, discourse ethics needs to recognize Hegel’s insight: that all formalism presupposes some context from which to abstract, together with material presuppositions about the self, the social institutions in which deliberation takes place, and so on. Justification must entail a degree of abstraction from context, but this abstraction cannot ever be cut sharply from the context, otherwise the completely decontextualized terms lose all meaning and the conclusions lose all possibility of application.

According to Benhabib (*op. cit.*, 42), the answer lies in the “comprehensive reflexivity” that discourse ethics entails. Nothing is beyond the scope of discourse, not even the presuppositions of discourse themselves. In everyday communicative action, we “practice the reversibility of perspectives implicit in adult human relationships,” and in discourse we “extend... this to the viewpoint of humanity” (52). The expression by others of their own viewpoints develops, Vygotskian-style, our wherewithal to take up multiple stances, provided we are open enough to attempt to empathize. This attempt, jointly and symbiotically, to enlarge our mentality enables us to address questions of the justification of norms, but without becoming unencumbered selves (73). We come into the conversation as ourselves and continually check any emerging positions against our lifeworld expectations (and vice

versa) in a process of reflective equilibrium. In such a way, we are able to achieve “reflexive role-distance” from our unreflective selves (73), without cutting ourselves free from them.

While Benhabib’s account shows how we can seek agreement without abstracting away from our own lifeworld, it does not fully address the problem of the alleged vacuity of abstract norms. To use one of my previous examples, this problem would arise if, although all the participants in a discourse have come to consensus that “innocent people ought not to be harmed unnecessarily,” the contents that different participants have for key concepts such as “innocent people,” “harm,” and “unnecessary” are so different that the assertion as a whole lacks any common content.

My claim is that, even though all proposed norms, including highly abstract ones, are uninterpretable in the absence of links to the lifeworld, yet human commonality will anchor a core shared content for these norms. This shared content may not be large, but it is sufficient to save abstract norms from vacuity. I will return to Johnson’s work, specifically his discussion (1993, 90-99) of the meaning of moral concepts, to explicate why all such concepts are linked to the lifeworld. He claims that much of the confusion about moral concepts arises because concepts are conceived of as being defined in terms of necessary and sufficient conditions. If concepts were of such a character, then it would be at least possible that continued discourse would be able to define them tightly in a way that leads to consensus on their meaning. Equipped with such tight concepts, discourse could then disconnect the concepts from the lifeworld and consider them in the abstract. Drawing on the pioneering work of Rosch (e.g. Rosch and Lloyd 1978), Johnson claims that concepts have an entirely different structure: they are prototypical.

Each person learns a concept by means of generalizing particular instances. Johnson uses the concept “lie” as his example. In a speech community, core instances of lies are so widely agreed that they appear to be structured by necessary and sufficient conditions - the speaker believes the statement to be false; the speaker intends to deceive; the statement is false - which are weighted in importance in that order. Depending on many factors, such as the usage of the word “lie” by those about one, one’s detailed history of incidences of lying (together with the outcomes) and so on, different non-central speech acts become more-or-less lies. This radial structure to the concept not only means that the concept has fuzzy edges, but also that the details of the fuzziness vary from speaker to speaker.

The consequent indeterminacy of meaning implies that, in discourse, it cannot be assumed that concepts carry the same meaning for all speakers, and that any

negotiation about the meaning to be assigned to the key terms of any disputed norms must always be tied to consideration of the lifeworld. However, it is equally important to note that the shared core meaning of moral concepts entails that the indeterminacy of concepts is not a radical incommensurability. This “highly stable and determinate character of the prototypical cases” (*op. cit.*, 101) is sufficient to underwrite a good deal of agreement in many cases where moral norms are submitted to discourse. In abstract statements of possible norms, agreement can be reached when all participants have in mind this stable core for the concepts involved. When it comes to thinking about detailed application of the norm, however, the radial indeterminacy of the concepts means (to return to the earlier example) that just exactly what counts as an innocent person in specific difficult cases will remain contested.

Within a particular culture, the stable and determinate core of a prototypical concept is shared amongst members of that culture. If we are engaged in a discourse across cultures, however, we cannot be so sure of the commonality of the prototypical core to a given concept. This is why the level of abstraction of normative claims submitted to cross-cultural discourses needs to be sufficient to tap into the human universals identified by Brown (1991), and listed above. Statements of universal rights (such as the innocent people statement) generally make reference to human universals. To be an innocent person means that one was not responsible for a wrong. Since all peoples have a concept of right and wrong, and one of responsibility, there is a non-empty core meaning to the phrase “innocent person,” even though its extension to specific concrete cases may be contentious. Instead of applying to the intersubjective lifeworld of a particular culture (a very thick lifeworld), it applies to the rather thinner intersubjective lifeworld of humanity that Brown has described. This is sufficient to anchor practical discourse. When Habermas (1992, 138) argues that “although they may be interpreted in various ways and applied according to different criteria, concepts like truth, rationality, or justification play the *same* grammatical role in *every* linguistic community,” we may take it that he is pointing to this stable core of meaning that enables us to understand and use such concepts, no matter what language or community they are used in.

In conclusion, in this subsection I have argued that the scope-widening effect of discourse ethics enables us to distinguish between goodish claims and rightish claims along a continuum of potentiality for agreement. Claims that lie towards the rightish end are those that are capable of garnering greater agreement, striving towards and approximating universalization, without ever reaching it. It is this that marks out justice, even though its boundary with realms containing other, less agreement susceptible claims is not sharp. Our approach towards universalizability (as in the

attempt to reach a wider consensus across cultures and time) must be tentative. Some moral contentions gather wide agreement (notably within the modern West, but also to some extent across cultures), and these tend to coincide with the norms picked out by Kantian theories as universalizable norms of the right and the just.

Justice is a virtue - a very important one. Yet, because agreement is neither unanimous nor certain, justice does not hold a unique, trumping position. It is one virtue amongst many, and since the virtues are not unitary, it may urge us to action different from that urged by other virtues. Discovering the best thing to do must always take place in contextualized deliberation. In the face of our knowledge of the potential fallibility of our concepts of each of the virtues, including justice, we need to reflectively balance the requirements of the different virtues whenever they seem to point us in different directions. Justice is an important touchstone, but not the only one.

7.25 Virtue in Discourse

Justice is not the only virtue, I have just claimed. What, then, of the other virtues, and how are they involved in a virtue ethics that draws on the insights of discourse ethics? I have already outlined the role of two broad categories of the virtues.

Reasonableness, we have seen, can be characterized in terms of the intellectual virtues. In this, I agree with Aristotle's view that "philosophical wisdom and understanding and practical wisdom [are the] intellectual [virtues]" (1980, I.13, 1103a5). Indeed, the multi-aspectual nature of reasonableness allows us to explain the dispositional character of the intellectual virtues more easily than a more rationalistic account does. This is due to the committed aspect to reasonableness, which is to some degree always present in any reasonable thinking. In order for us to be able to display (for example) the disposition to be intellectually honest, or to manifest intellectual courage, we have to be committed to our thinking (§3.1236) .

The second role for virtue that has already been outlined is the ethical virtue (or bundle of virtues) called "respect." In §6.24 and elsewhere, I have argued against the Habermasian view that discourse ethics rests on a purely formal principle. Instead, I have asserted that it requires a basal value: respect. In Aristotle's listing of the virtues (ibid., II.7), respect does not have a high profile. Of the ten virtues he discusses, three are to do with the "intercourse of words and actions" (ibid., 1108a10); that is, with the relationships between persons in discourse and interaction. The names he gives the means of these three virtues are "truthfulness," being "ready-witted" and "friendliness." Certainly, we can see that each of the three is implicated to some extent in respect for our interlocutors, but it is not hard to see that a full account of respect

(which, of course, I have not offered) would include a number of other more central virtues of respect, such as care for their feelings, taking what others say at face value in the absence of contrary evidence, hearing them out and so on. Here it is obvious that the dialogical (rather than monological) core of discourse ethics alerts us to just how important the virtues of respect are in human lives.

To refer to the intellectual virtues and the basal virtue of respect, however, is not to exhaust the relationships between discourse ethics and virtue. Discourse ethics, we recall, is based on Habermas' account of the place of values and norms in our lives. Ordinarily, he claims, we operate within the normative horizon of our society, but when one of the norms that make up that horizon becomes problematized, we suspend action, enter practical discourse and try to redeem the norm. To recast this in virtue ethics terms, we usually act habitually and virtuously until the virtuousness of a particular action is brought into question. At that time, we bring practical wisdom to bear to determine the right way to act. Here we see the difference between discourse ethics and virtue ethics. Within discourse ethics, the primary mode of inquiry into normative questions is dialogical (and individual reflection is parasitic on that), whereas for a virtue ethics, the order is reversed.

Given my developmental interest, the difference is vital. In moral education, in the light of the Vygotskian account of learning, it is clear that discourse ethics provides an essential modification to virtue ethics. As I shall explore more thoroughly in §9.2 and §10.23, moral education needs a mechanism whereby the virtues become the subject matter of a practical discourse. I shall suggest that the community of ethical inquiry provides that mechanism.

Virtues do not form merely the subject matter of discourse. If we are to achieve the aims of moral education, then it is not enough that students consider the virtues: they must also become virtuous. A detached and "rationalistic" discourse on virtue which did not lead the participants to become more virtuous would not be contributing to moral development. Aristotle warns us that mere philosophizing on its own is insufficient for becoming virtuous: "But most people... take refuge in theory and think they are being philosophers and will become good in this way" (ibid., II.4, 1105b13). Hence, the specific characteristics of the practical discourses that students enter must contribute to a linking of the conclusions of the discussion to the formation of virtuous habits. Because the details of the dialogue are so important, I will not address this issue just yet: I am presently talking at too high a level of abstraction. I will return to the issue in §9.2, where my focus will be on classroom organization.

7.3 From Meta-ethics to the Classroom

Chapter Seven has gathered together some of the major strengths and weaknesses of virtue ethics and discourse ethics as meta-ethical systems. This survey has, I believe, shown why virtue ethics provides a suitable account of morality to use as a basis for a moral education program. Virtue ethics has the advantages that it emphasizes the development of character as a multifaceted and idiosyncratic achievement by encumbered individuals within a society. However, classical virtue ethics, developed as it was in a relatively closed society, needs supplementing with some sort of recursive and fallible, yet still useful, guidance as to what constitutes moral progress in a multicultural and pluralistic society. Further, virtue ethics neither provides more than a sketchy outline of the place of education in moral development, nor does it consider sufficiently the role of discussion in such development.

Discourse ethics offers resources for escaping the conformist and traditionalist dangers of virtue ethics. Moral improvement for society as a whole is not to be found in the unreflective following of the precepts of other, wiser folk. Nor it is to be found in individual soul searching, unless this is parasitic on the quest for intersubjectivity. As an account that takes cross-cultural discourse seriously, and provides considerable insight into its educational importance, I turned to a version of discourse ethics shorn of its Kantian transcendentalism. While I have argued that Habermasian discourse ethics fails as a complete meta-ethical system, there is much to be gained from its insights in educational terms.

In terms of the primary aim of this thesis, discourse ethics can play a vitally important role: providing a richer and more practically useful account of the development of moral character. There are, as I have pointed out above, distinct parallels between the virtue ethics mix of habituation with reflection, and the discourse ethics mix of socialization in the lifeworld with discursive exploration of norms. Discourse ethics, however, supplemented by both a broadening of rationality into reasonableness, and a reconstruction of its developmental model to take in Vygotskian as well as Piagetian insights, can give us a much clearer insight into the role of discussion in school-based moral education.

To give an account of the place of the community of ethical inquiry in a moral education program is to be the task of Part III of the thesis. Before I proceed to that, I will take stock of what I have done so far. Parts I and II of this thesis have laid the major philosophical basis for a justification of the place of a Community of Inquiry in moral education. Part III will draw on these conclusions in the context of the classroom.

Any account of moral development must take notice of the fact that we are talking about the moral development of human beings, who come into the world with the features characteristic of that species. In Chapter Two, I looked at the biological roots of moral development, showing that, while human biology cannot be said to fix the outcome, it needs to be taken into account, and provides some commonality.

In Chapter Three, I investigated one of the core capabilities essential to moral agency, often referred to as rationality or reason. I argued (§3.1) that a full account of this capability requires attention to five aspects - the critical, creative, committed, contextual and embodied aspects - and that we ought to use the name “reasonableness” so as to emphasise that it is a much broader capacity than has been commonly picked out by “rationality” or “reason.” Reasonableness so construed is capable of describing the intellectual capabilities of real persons whose thinking takes place while they are immersed in the physical and social world. I also argued (§3.23) that an understanding of the mechanisms by means of which reasonableness develops is also vital to a proper understanding of reasonableness itself, and examined the Vygotskian explanation. Reasonableness must be learned through interactions in the social world, scaffolded by more experienced thinkers. Such a recognition leads to a reconceptualization of thinking for oneself (§3.241) and the realization that autonomy requires interdependence, not separation (§3.3).

A focus on Habermas’ account of communicative action in Chapter Four provided a theoretical framework within which to explore more closely the development of reasonableness and normative judgement within the classroom. Initially, I explored the role of strategic action, communicative action, critical discourse and the ideal speech situation in Habermas’ thought (§4.12), before turning to an exploration of how these ideas can assist in the analysis of classroom dialogical action (§4.13 and §4.14). I argued that, in order to understand how teachers can assist students in developing reasonableness and autonomy, we need to modify Habermas’ categories of action to incorporate pedagogical action: an inextricable mix of strategic with communicative action undertaken by a more experienced person with the less experienced, so as to lead to such development. Reconsidering autonomy in §4.2, I advanced the case that autonomy must be reconsidered as communicative autonomy, whereby children are equipped to participate effectively in critical discourse.

I turned my attention to moral theory in Part II. In §5.1, I argued that any meta-ethical theory that takes development seriously needs to be multi-dimensional, rather than basing morality on a single principle. Aristotle’s virtue ethics is such an account, and I explored it in §5.2, arguing that the traditional interpretation that Aristotle requires that children be mechanically habituated into morally good habits before they can

engage in philosophical and ethical discourse is mistaken. On the contrary, I argue that habituation requires judgement, and that judgement can be built through the engagement in philosophical inquiry.

For a better account of the place of dialogue in ethics, I turned in Chapter Six to Habermas' discourse ethics. Firstly, I outlined Habermas' project in §6.1, before appraising the program in §6.2. I argued that the weaknesses in discourse ethics - its assumption that ideal discourse will (in matters to do with the right) inevitably reach consensus; its inadequate account of rationality; its claim to be based on a purely formal principle - mean that it is not satisfactory as a complete meta-ethical system. Nevertheless, when reconceptualized to cover these weaknesses, it has much to offer.

In Chapter Seven, I looked (§7.1) at the ways in which virtue ethics can be strengthened by incorporating many of the insights of discourse ethics, particularly in terms of the necessity of taking respect as a foundational virtue (§7.21), the place of dialogue in development (§7.22), a way of allowing for progress within virtue ethics (§7.23) and an account of justice as a virtue (§7.21).

In this summary, I have gathered together the major ideas that will inform the three chapters of Part III. It is time to focus much more directly on the classroom, the students in it and the teacher.

Part III: Discourse and ethics in the classroom

Chapter Eight: Ethics, the School and the Community of Inquiry

In Parts I and II, I have been concerned to look philosophically at a number of concepts which are centrally important to moral education, especially when conducted through discussion of ethics in the classroom, as summarized in the final section of Part II. Chief amongst these are reasonableness, autonomy, pedagogical action and two meta-ethical systems: virtue ethics and discourse ethics. In Part III, I will draw together these key concepts specifically with an eye to what teachers can do in their own classrooms in order to assist children in the task of growing up to be ethical persons.

In this first chapter of Part III, I will acknowledge that teachers can play many roles in such assistance, before concentrating on a positive move that they can make: the engagement of children in a community of ethical inquiry. The second section of this chapter will investigate the notions of community, inquiry and why their intersection in a community of inquiry can be expected to be educationally powerful.

In Chapter Nine, I turn my attention to a number of the bigger issues identified so far, and outline just how they can be addressed within the community of ethical inquiry. Firstly, I will show that claims that the community of inquiry is an efficacious medium for the development of reasonableness are supported by empirical studies, both of other critical thinking programs, and of Philosophy for Children itself. The second section will explore the place of the community of ethical inquiry in ethical development, addressing two claims of weakness that are sometimes levelled at classroom-based moral education programs: that they cannot lead from judgement to morally good action; and that they produce “bush lawyers” who can better rationalize their wrong actions. In the third section, the place of the teacher as a wielder of pedagogic power and action is taken up again, with specific attention to the community of inquiry. Finally, I will look at the charge that a teacher who wields pedagogic power must be indoctrinatory, and also consider the possible consequences of a retreat from pedagogical action: ethical relativism and moral indifference.

The focus of my attention in Chapter Ten will become even more practical. There, I will look at the specific stages of the community of ethical inquiry and discuss specific steps that teachers can take in best conducting these communities. I will look in turn at the questions of the ways of setting up discussion, the control of the content of the discussion, and the inculcation of reasonableness, moral correctness and

communicative autonomy. In each case, I will show what the philosophical underpinnings are for my advice, drawing on Parts I and II.

The key claims to be established in Part III are that the task of moral education is to assist children to develop their reasonableness (i.e. their intellectual virtues), their ethical and moral virtues, and hence their autonomy. One powerful method of achieving these aims (which are interrelated in ways to be explored) is to engage children in ongoing ethics-oriented communities of inquiry which are run by teachers wielding pedagogic action.

8.1 Moral Education and Schools

I use the word “assist” in the last paragraph because it ought to be obvious that teachers neither should nor could be the sole source of moral education, let alone morality, for children. Firstly, many other people and institutions, in particular members of the close family, have roles that impinge on moral development. In the case of the family and probably some others, this influence is far beyond that which a single teacher can have. Secondly, the account that I have been developing implies that the children themselves ought to be taking an increasing responsibility for their own moral development. It is not something that can be imposed on a child from the outside.

Nevertheless, given the focus of this thesis, I shall have little to say about these matters, apart from pointing out that these other influences are not always benign, and the teacher’s role may include what we might call remedial work on damage done by families, the media or other agents. This, of course, raises disturbing questions about the rights or duties of teachers to engage in educational activities that might work against the views of families or broader society. The questions of indoctrination (§9.4) and the place of the teacher’s personal moral views (§10.23) will be addressed subsequently.

Some make a sharp distinction between moral education and moral training (Beck, Crittenden et al. 1971, Introduction, 8-9). In that collection, for example, Baier (1971, 95) claims that moral education must be both deliberate and carried out by professionals with the aim of fostering moral excellence: Kohlberg (1971) takes a similar view. Like others in the same collection (Gauthier 1971; Loubser 1971; Melden 1971), I will not make a sharp distinction. My discussion of the Aristotelian notion of reflective habituation (§5.23), when bolstered by a Habermasian account of the necessarily communicative element in such reflection (§6.22), undercuts any such analysis. Parents or others who offer even the most simple of reasons for their moral directions to children are engaged in moral education just as much as moral training.

In this chapter, I will follow a path that has been hinted at in §6.24. While there are, of course, many interesting questions to address about cross-cultural issues in moral education, I shall restrict my focus here to moral education in modern Western societies. This is not to say that such a restriction will make the issue of cultural diversity disappear. On the contrary, given that modern Western societies are culturally diverse, I will need to pay considerable attention to this issue. Rather, it is to say that I will take the basic underlying moral principle of respect as given. I argued in §6.24 that respect must be assumed in order to underpin discourse ethics, and I take it to be one of the defining moral notions of modern Western moral thought. Consequently, when considering moral diversity and the response of schools, I will assume that schools, being institutions of a modern Western society, ought to take a conception of respect as given. However, just what this concept of respect constitutes, in detail, and what practical implications it has will remain contested (see 10.23). Thus it will be one of a number of key concepts that ought to form the subject matter of classroom discourses.

For the moment I will concentrate on the school itself. A school is an institution set up with the broad aim of educating children. Just how we are to understand this broad aim itself is, of course, the subject of much philosophical work, ranging back to Plato's *Republic* and beyond. Whether education ought even to include moral education has been questioned by some, particularly those who have what Dewey (1966, 359) calls "a narrow and moralistic view of morals." On the one hand, some of these assert that schools ought not to have anything to do with moral education because any attempt to inculcate a moral view amounts to indoctrination or the enforcement of religion (§9.4). Others assert that moral education is the domain of the parents, not the state, and schools should keep out of it. I shall argue that schools cannot withdraw from moral education even if they attempt to, for the stance of avoiding addressing morals inevitably involves a transmission of certain precepts with moral import.

8.11 The Range of School Procedures with Ethical Import

Teachers and schools are important. Whether teachers like it or not, even accept it or not, they are inevitably influential in the moral development of the child. Much of this influence is incidental (which is not to say it is beyond the control of the teacher). It occurs through the model of moral concern and action that the teacher presents, through the expectations the teacher has of children in the class, through the rules of the classroom, through the habits of behaviour that the teacher inculcates.

Nor does the teacher exist in a vacuum. The school as an institution has its own impact on the moral development of the children through many avenues, as diverse as the school rules; the model it presents through its discipline policies; the informal interactions that the school encourages, accepts, ignores and discourages; the formal messages conveyed at assemblies and other occasions of public messages; the interactions it has with the wider community and so on.

I am not making a claim that any of these types of practice are, at least in general terms, bad or unnecessary. Most of them are inevitable, and each can be implemented well or badly - that is, to assist or to hinder the development of ethical persons.

It is not my intention to develop a typology of all the ways in which a school influences the moral development of the children entrusted to its care. Nor do I intend to analyse each influence in any detail. Many of these influences are not at all, or only partly, explicit. Others are fully explicit, but are in the form of advice or orders from on high. Each of them can be, and not uncommonly is, made the subject of explicit evaluation within a school, as the school attempts to do its job well.

By and large, however, these influences of the school on its students can be put into one of two categories (though there may well be considerable fuzziness at the boundary between these categories): the influences of the school fall either into the category of relatively unreflective moral shaping, or the category of moral instruction. In the category of unreflective moral shaping, I place all those practices which require students to act in certain ways, or model such action to students, with little or no explicit reference to the moral import of the actions involved. Moral instruction does make such explicit reference, and it includes what might be called sermons or lectures: moral influence purely through one-way talk.

A few comments on this categorization are needed. Firstly, the use of the word "moral" here is neutral as to whether the influence is morally praiseworthy or blameworthy. Both unreflective moral shaping and moral instruction can have bad as well as good moral effects - and sometimes it will be contentious as to which it is. Secondly, it is not hard to see why these categories overlap, for the explicitness of moral referencing can vary smoothly from none at all to quite up-front. Thirdly, it is quite conceivable that a single incident may have mixed moral messages, such as when a teacher shouts angrily at a child, "You shouldn't try to scare others into doing what you want them to do!"

Those who advocate that schools ought not to be involved in moral education generally have in mind only the category of moral instruction. They ignore the fact

that much of the moral influence of a school lies in the unreflective moral shaping - part of what has been dubbed the “hidden curriculum.” A school which treats children respectfully teaches the value of respect and produces children who are more respectful of others, whether it explicitly preaches respect or not. In a school where teachers are explicitly banned from expressing any moral view, students would be likely to learn the moral lesson that to have a moral stance is not important, or that any moral stance whatsoever is acceptable. Of course, it is doubtful that such a school could exist - it is usually *contentious* moral views concerning matters in broader society that are being discouraged.¹ Those moral views that are taken-for-granted within the lifeworld are not only largely invisible to unreflective participants (as we all are most of the time), but are also continually reinforced. Any time teachers insist that their class is quiet when they are speaking, they are tacitly expressing the view that respect is due to them. If they similarly insist on quiet when another child is talking to the class, the message is extended to respect for peers. Indeed, any action teachers take with regard to students’ behaviour will have some moral import, even if very minor.

As Habermas makes clear, any of the taken-for-granted norms of the lifeworld can be problematized by reflective participants at any time (§6.1). Hence, any of the moral influences on students mentioned above may potentially be made the subject of a practical discourse. Schools and teachers that are concerned about their impact on their students engage in investigations into these matters, and educational researchers (especially of an ethnographic bent) mount studies to discover them. It is of course also possible, though it usually happens much less frequently, to include the students or their representatives in such practical discourses, either through such institutionalized bodies as student representative councils, or in other more informal ways.

Of course, when schools, teachers and/or students enter into discourse about moral matters, there are many different forms that the discourse can take. As I have commented in respect to student involvement, some are highly regulated and institutionalized, others are quite informal. We have seen in §4.124 that Habermas’ ideal speech situation can play a regulative role in assessing the adequacy of discursive set-ups. The rules of institutionalized discourses can serve to maintain the power of some and to exclude others, or they can serve to neutralize power differentials and to be inclusive. Similar comments apply to informal interactions, even though the rules will be both more tacit and more fluid in such circumstances, and will depend on the

¹ I will note in passing here that the opposite view - that teachers ought to express their views on contentious moral issues - also has its dangers. I shall return to this issue in §9.4.

persons involved and the details of how they interact with each other. Further, the informal nature of this sort of discourse may serve to make it harder to ensure that distortions and exclusions are either discoverable or preventable.

Thus if we are to recommend that practical discourse is a vital part of moral education, then it will be necessary to have an account of how such discourses are to be set up, and what sorts of rules and understandings are to structure it. In the remainder of this chapter, I will argue that the community of inquiry is a suitable structure for real practical discourses. Since the phrase “community of inquiry” has been used in a number of related but differing senses, I will explore these usages, before concentrating on developing a version suitable for classroom practical discourse, and explicating the reasons why a classroom community of ethical inquiry ought to be seen as an essential element of moral education.

8.2 The Community of Inquiry

The origin of the phrase “community of inquiry” has been (tentatively) traced back by Lipman to Peirce’s article *The Fixation of Belief* (1955, 5-22). This is what Lipman has to say:

This phrase, presumably coined by Charles Sanders Peirce, was originally restricted to the practitioners of scientific inquiry, all of whom could be considered to form a community in that they were similarly dedicated to the use of like procedures in pursuit of identical goals. Since Peirce, however, the phrase has been broadened to include any kind of inquiry, whether scientific or nonscientific. (1991, 15)

Peirce certainly had such an extension in mind, for he also talks of inquiry amongst the community of philosophers (1955, 229). However, Peirce does not say a lot about exactly what a community of inquiry is. Lipman, and (following him) others in the Philosophy for Children movement, have written extensively on the notion, but usually in the context of education. There exists a tension in this writing between expansive and particular notions of the community of inquiry, which I shall explore in Chapter Nine. Before I turn to that, though, I shall explicate the notion of a community of inquiry in the broader sense of a community existing throughout society, as indicated by Peirce.

Clearly, an obvious way of analysing the notion of a community of inquiry would involve some sort of characterization of “community,” followed by an examination of “inquiry,” and finally an explication of the way in which inquiry applies to community. This is the path that I shall follow. In this analysis, I will restrict my focus to human communities.

8.21 Community

The first thing that is obvious about a community is that it must consist of a collection of individuals. But a community is more than just any group of people. As the etymology of the word itself indicates, a community must have something in common. Now, there are many ways in which a group of people can have something in common, but not all of them turn the group into a community. Originally, the commonality was a quasi-political one, so that a community lived in close proximity and was subject to the same laws and regulations. The term is still used widely in this sense. By extension, we now also talk of communities as groups of people who have similar aims, purposes, intentions and/or beliefs.

In discussing the concept of a community, it is useful to keep Johnson's point (§7.24) that concepts are not generally defined by necessary and sufficient conditions in mind. There are two ways we can apply it here. The first applies to the totality of different communities. They neither all necessarily have the same type of thing in common, nor is it sufficient for a group of people to have some common factor to form a community. Equally, it is not the case that we can say that a group of people either are a community or they are not. Some groups, with similarities and links that place the group close to the prototypical core of the concept, are clearly communities. An example might be an isolated village in the Highlands of Papua New Guinea. Other groups form only, at best, loose communities, such as the community of players of a particular computer game. Many of them do not know many other members of the group, and would not be too bothered about communicating with them. Yet they do have an interest in common, and if they do meet, are likely to be interested in discussing this common interest, perhaps by swapping tips for playing it. Some groups of people that share something in common - say, a mole behind their left ear - cannot be said to form a community in any but the most vague and metaphorical sense.

Secondly, the boundaries of a given community are also often far from clear. Some communities have membership lists, explicit rules, common commitment to a cause and other trappings of institutionalization. These tight communities are those that help define the prototype for the concept community. It is relatively easy to identify members of the community, as well as criteria for joining and leaving it. Yet even in these tight communities, there can be demarcation disputes. Is a member who has forgotten to pay the membership subscription part of the community or not? What about the interested newcomer who is attending meetings but has yet to officially join, or the long time member who still pays, but has not attended meetings in years and hardly ever thinks about the organization? Considerations like these show that a community is not coextensive with an organization that formalizes the community.

And if there are boundary problems with communities based around organizational structures, these become much more acute when dealing with less formal communities, such as Peirce's communities of scientists and of philosophers.

As the examples of the potential and the inactive members of a community show, a community is not a static entity. Individuals can become more involved in a community over time, or less involved. As they do so, the community changes. The community can become stronger, and thus closer to the core of the prototype, or it can become weaker and possibly disintegrate.

In modern societies, as Dewey (1966, 20-22) reminds us, communities are internally diverse, and individuals are members of many overlapping communities. "Each such group exercises a formative influence on the active dispositions of its members," so our membership of communities is partly constitutive of us, just as the nature of the community itself is influenced by the characteristics of the people who join it. In part, it is the diversity of these formative influences present in any large community (such as a modern Western state) that acts against the production of a stifling conformity.

This brief survey of the concept of a community leaves us with a number of important points to consider when we turn our attention to the particular nature of a community of inquiry in Chapter Ten. We shall need to identify the commonalities that distinguish the community of inquiry from a mere group of disparate individuals, all the time being aware that these commonalities may not be shared by every member. Secondly, we shall need to pay attention to the ways in which individuals become members of the community, or maybe drop out of it, and how these processes help the community to grow and evolve, or possibly decay. Thirdly, we will have to be aware of the diversity within the community, including both the possible existence of sub-communities and the overlapping of the community with many other communities. Finally, we will need to consider how the community interfaces with other, larger communities of which it is a sub-community.

8.22 Inquiry

Peirce (1955, 10) defines inquiry in this way:

The irritation of doubt causes a struggle to attain a state of belief. I shall term this struggle *Inquiry*... With the doubt, therefore, the struggle begins, and with the cessation of doubt it ends. Hence, the sole object of inquiry is the settlement of opinion.

More recently, Splitter (1997, 30) has this to say about inquiry:

Inquiry... what we have to do when we want to understand the why and wherefore of things. An inquiry is, fundamentally, a search: a search for answers to questions and

puzzles that confront us - or better, a search for understanding (which may lead to answers, but does not even assume that we know what questions to ask).... Inquiry is, necessarily, self-correcting. This reflective - or reflexive - dimension of inquiry reminds us that inquiry is a mode (or perhaps, many modes) of thinking.

It is certainly worth noticing the similarities: in both, inquiry begins in doubt, questions, puzzles; it continues in a struggle or search and it ends in the settlement of opinions, or answers. But then Splitter goes on to make an interesting modification to this apparent unanimity. He suggests that inquiry aims beyond merely coming to answers (the settlement of opinion), towards a more encompassing outcome: understanding. As he points out, this is more encompassing because the search for understanding may or may not issue in answers. One can see the roots of Habermas' ideal consensus (§6.21) in Peirce's conception of inquiry.

Splitter's characterization may be wider than Peirce's in another way. Can we so easily identify "the irritation of doubt" with "questions and puzzles" as I did above? Again, Peirce is anticipating Habermas, who claims the trigger for discourse is the calling into question of the taken-for-granted of the lifeworld. When irritated by doubt, or when our taken-for-granted knowledge is problematized, we need to come to a resolution. Yet, although questions and puzzles can arise in this way, they may also arise from curiosity rather than need.² My juxtaposition of curiosity and need may seem to call for further analysis, for there does seem to be a sense in which curiosity creates a need - to find out - of its own. Nevertheless, given Peirce's reference to the *irritation* of doubt, I do think that we can contrast the drivenness of this type of need with the playfulness of curiosity. Further, the explanation of Peirce's favourite model for inquiry - science - would be sadly lacking if inquiry arose solely from the irritation of doubt, and left no room for that well-spring of pure research: the imaginative asking of "What if..." questions. Curiosity fits better if we have a more expansive conception of reasonableness, with full recognition of the creative aspect of thinking, rather than a narrow, critical thinking based idea of rationality.

Splitter's final claim is that inquiry must be self-correcting. Here I think that we would find that Peirce and Habermas would agree, and the reason why they would is instructive in our analysis of inquiry. As we have seen, both Peirce and Habermas lay a great deal of emphasis on the place, in inquiry, of community. They both rule out the possibility of substantive gains from purely monological thinking. Even if an idea can be introduced into discourse from a monological position, it must be tested in community before it can have any claim to legitimacy (Peirce 1955, 13; Habermas

² I must thank Cyn Townley (pers. comm.) for drawing to my attention to the roots of inquiry in curiosity as well as need.

1990, 66). It is this public testing of ideas that ensures reflexivity. In the discourse in which reasons are offered for an assertion, counter reasons may be advanced, and any flaws in the reasoning may be exposed. Indeed, because we learn the processes of reasoning and inquiry within such public discourses (§3.23), then even that part of inquiry that can be carried on individually will be subjected to the internalized voices of critique and reflection (Lipman 1991, 52). Here we begin to see the reason why “community” and “inquiry” merge so well into the community of inquiry.

8.23 The Community of Inquiry as a Setting for Vygotskian Learning

Underlying the notion of a community of inquiry is a concept which I refer to as Vygotskian learning and outlined in detail in §3.23. In this section, I shall recall what I mean by Vygotskian learning, and draw together the strands of it that have been variously developed in different places throughout this thesis. Then I shall study more closely why it is that the community of inquiry offers an exemplary setting for Vygotskian learning.

According to Vygotsky, we recall from §3.23, “What the child can do in cooperation today, he can do alone tomorrow” (1962, 104). This is the key Vygotskian insight: that children can engage verbally in a social group in forms of reasoning that are not yet available to them individually, and that it is precisely this engagement that enables children eventually to use those forms of thought. We must remember that the phrase “forms of thought” must be construed quite widely, as befits both the broad version of reasonableness developed in §3.1, and Vygotsky’s own broad conception of thought (§3.234). I have argued, for example, that the less reflective versions of the virtues can be learned Vygotskian-style (§5.232), as well as practical reasoning itself.

Vygotsky uses the term “zone of proximal development” to label this conceptual and reasoning space in which children can operate with help from a group, but are not capable of operating in on their own. The community of inquiry offers an excellent environment for a group of students to construct such reasoning, especially with the scaffolding provided by a teacher engaged in pedagogical action (§4.141).

Lipman (1991, 52), the foremost theorist of the community of inquiry in school education, provides extended examples of the sorts of capacities that can be internalized in this way.

Characteristic behaviours of the community	Internalised individual behaviours
Members question one another	Individuals question themselves
Members request of each other reasons for beliefs	Individuals reflect on their reasons for thinking as they do
Members build on one another's ideas	Individuals build on their own ideas
Members deliberate among themselves	Individuals deliberate in their own thinking
Members offer counterexamples to the hypotheses of others	Individuals anticipate counterexamples to their own hypotheses
Members point out possible consequences of one another's ideas	Individuals anticipate possible consequences of their own ideas
Members utilize specific criteria when making judgements	Individuals use specific criteria when making judgements
Members cooperate in the development of rational problem-solving techniques	Individuals follow rational procedures in dealing with their own problems
Internalization of social thinking (after Lipman, 1991)	

It is worth noting that this Vygotskian internalization of social thinking can be both relatively unreflective and explicitly reflexive (or somewhere on the continuum between them). Those less experienced persons who engage in the social group may well pick up competency without ever bringing the matter to focal attention - an unreflective habituation to an intellectual virtue. On the other hand, especially if the more experienced members of the group (in the classroom case, particularly the teacher) flag the possibility, the competencies being learned may become the explicit content of discussion. A turn to Habermasian analysis (§4.1) will make the connections between Vygotsky and my broader argument clearer. Unreflective appropriation of situated rationality (§6.23; i.e. my "reasonableness") is the norm in the lifeworld. Communicative and, especially, strategic action will commonly result in unreflective learning, while a shift to critical discourse (§4.123) marks the suspension of lifeworld action, and the reflexive attempt to understand (amongst other possible topics) the competencies of the lifeworld.

In the community of inquiry, both processes can take place. While the students are driving the inquiry, many of the competencies in the right hand column of the list above will be internalized relatively unthinkingly. Yet the teacher, will not be standing idly by. By engaging in pedagogical action (§4.141), the teacher can both encourage this process through the strategic element of pedagogic action, and help the students (when appropriate) to turn their attention to the competencies of reasonableness through the communicative and, especially, discursive elements. The latter role of the teacher in scaffolding explicit rather than just implicit learning of reasonableness is one

of the features that marks off the community of inquiry from normal lifeworld dialogue. It is also the way that students become progressively more able to think for themselves (§3.241) and hence intersubjectively work towards autonomy (§3.3).

In an excellent comparison of the Piagetian and Vygotskian programs, Howe (1996) considers their implications for science teaching. While the focus of my dissertation is moral rather than science education, Howe's discussion raises a number of relevant issues. The implications of Piaget's program are that the teacher needs to create a classroom climate in which children construct and test their own hypotheses, in interaction with the world. In doing so, the teacher creates the setting for children individually to advance a further stage through their own efforts. This is, in Piaget's terminology, a form of constructivism: children construct their cognitive capacities through individual interaction with the world, and hence must be placed in situations of cognitive conflict.

Howe asserts that the implications of the Vygotskian program are different: teachers need to treat children "as participants in a joint enterprise in which meaning is derived through interaction with other people, mediated through language" (ibid). Language must be treated as a tool to support and promote thinking, with the teacher taking a central role of stimulating children to reflect and explain. A Vygotskian perspective emphasizes the importance of context and questions the automatic transfer of broad skills. Howe claims that the constructivist insights of Piaget need to be merged with sociocultural views, encouraging the integration of personal experience and knowledge into the scientific systems, models and ways of thinking that are being learned. In particular, he calls for a shift of focus from the solitary thinker to the thinker/talker in a social context.

What is true of science teaching, which has an explicit focus on the physical world, is even more true of moral education. Ethics deals with contents which are social constructs (even though they may be underpinned by biological facts, as I argued in Chapter Two). This Vygotskian picture is supported strongly by a Habermasian account of progress towards moral correctness through engagement in a practical discourse (§6.22 and §7.23).

Empirical research provides support for the Vygotskian contention that the major vehicles of thought are constructed socially in dialogue. Based on Perkins' (Perkins 1994) analysis of open-ended dialogue, I investigated the structure of classroom talk in a community of inquiry (Sprod 1994c; 1997c, d; 1998: more details of this research can be found in §10.222). The analysis was able to identify episodes, made up of utterances by different participants in the dialogue, that followed a single strategy (or

epistemic game) for investigation of the ideas being discussed. These episodes, which I dubbed “epistemic episodes,” are the jointly constructed thinking strategies to which Vygotsky refers. Analysis of the lesson transcripts provided clear evidence that students became more likely to use, and better able to sustain, epistemic games that had been previously introduced and sustained by the teacher (Sprod 1998, Table 6).

Similar episodes have been independently identified in other discourse analysis projects. Bruner (1983) studied 8-13 month old children and their interactions with their mothers through the game of “peekaboo.” He found that, over this time, the children appropriated the deep structures that lay beneath the mother-child interactions, so that the child could eventually control the interaction in the same way that the mother initially had. He labelled these thinking structures “formats”, and comments

It is at the heart of any support system involving games - "play" games and language games alike. If the teacher in such a system were to have a motto, it would surely be "where before there was a spectator, let there now be a participant". One sets the game, provides a scaffold to ensure that the child's ineptitudes can be rescued or rectified by appropriate intervention, then removes the scaffold part by part as the reciprocal structure can stand on its own. (60)

Tizard and Hughes (1984) studies the interaction of somewhat older children - preschoolers - and identified “passages of intellectual search,” in which parents and children (but not nearly so often preschool teachers and children) engaged in episodes displaying similar characteristics to my epistemic episodes. These, they asserted, were influential in teaching higher order thinking to the children. Similar claims have been made by Barnes and Todd (1977), who labelled similar sequences of talk “episodes of cognitive strategies” in their study of small group work, and Pontecorvo and Girardet (1993), who in their account of research on small group work in history classrooms, used the rather unwieldy label “reasoning sequence which uses a single epistemic action.”

Thus, there is substantial empirical evidence that the very ability to think well, and the concepts with which to think, are both dependent on dialogical social interaction which seeks conclusions. While such interactions are a common part of the lifeworld, they are often informal and unfocused. This is not surprising when we consider that such interaction ostensibly serves many other purposes as well. Since cognitive, conceptual and ethical development are all key aims of schools, they need to create these conditions in a more explicit and controlled way. In the next two chapters, I will argue that the obvious classroom setting for maximising these opportunities is a classroom community of inquiry.

In the course of this chapter, I have outlined the features of communities and of inquiry. My claim, following Peirce and Lipman amongst others, is that the

constitution of a community of inquiry (i.e. a group whose commonality is the commitment to engage in inquiry) is a powerful learning methodology. To substantiate this claim, I have called on Vygotskian theory. The Vygotskian account is highly compatible with several of the aims I have argued for in Parts I and II of this thesis.

Vygotskian learning is the key way to develop reasonableness, and its reliance on a rich interpersonal setting clearly enhances the development of thinking that displays the multi-aspectual character I have identified for reasonableness. Virtue as I have characterized it also gels well with Vygotskian learning, especially in the light of the Habermasian gloss I outlined above. Not only can the virtues be learned relatively unreflectively through immersion in the lifeworld, but the wherewithal for reflection on that habituation (a central part of habituation in the virtues - §5.21) is also learned both through the Vygotskian implicit habituation and through explicit reflexivity. This practical reason (in Aristotle's terms) can be then applied to reflection on the acquisition of the moral virtues. Thus it can be seen that we need to draw on Aristotle's virtue ethics, Habermas' theory of communicative action and Vygotskian learning theory to show the power of the community of ethical inquiry in moral development.

Chapter Nine: The Community of Inquiry in the Classroom

In the previous chapter, I gave a broad brush outline of the notion of a community of inquiry. Yet it is still a rather nebulous notion. As we shall see, the idea has been applied in a number of social settings. Coming from the Philosophy for Children approach, I am primarily interested in the classroom setting. I have, in the final section to the last chapter, linked the notion to a number of philosophical approaches (reasonableness, Habermasian theory, virtue ethics, communicative autonomy) which do not figure in its historical development. In this introductory section, I will look more closely at those historical roots, and at its usage in Philosophy for Children. Subsequently, each of these major philosophical themes will re-enter the discussion, as I consider the place of the classroom community of inquiry in moral education.

Drawing explicitly on the work of Peirce and Vygotsky, the Philosophy for Children movement has adopted the phrase “community of inquiry” and applied it to the classroom. Lipman (1991, 16), founder of Philosophy for Children, says:

Thus we can now speak of “converting the classroom into a community of inquiry” in which students listen to one another with respect, build on one another’s ideas, challenge one another to supply reasons for otherwise unsupported opinions, assist each other in drawing inferences from what has been said, and seek to identify one another’s assumptions. A community of inquiry attempts to follow the inquiry where it leads rather than being penned in by the boundary lines of existing disciplines. A dialogue that tries to conform to logic, it moves forward like a boat tacking into the wind, but in the process its progress comes to resemble that of thinking itself. Consequently, when this process is internalised or introjected by the participants, they come to think in moves that resemble its procedures. They come to think as the process thinks. (16)

This passage clarifies the concept of a classroom community of inquiry somewhat, but it also raises a number of questions. Firstly the clarification: the Vygotskian ideas that underlie the community of inquiry (§8.23) are very evident in the final two sentences. Moreover, the active participation of the students in co-constructing the arguments is also clear. But there are a number of problematic or obscure aspects to the notion. Nowhere in this passage is the teacher mentioned. Yet, as anyone who spends any time in a classroom knows, the teacher has a vital role to play in any classroom. I have discussed this in some detail in §4.14. The notion of “following the inquiry where it leads” also needs further unpacking, and may seem problematic in two respects: How can an inquiry, as opposed to some or all of the participants in the inquiry, lead? How are the agreed curriculum materials to be taught under such a system? I shall address these questions in §10.21.

Clearly, before proceeding, I will need to have a more precise account of what I will mean by the phrase “a classroom community of inquiry” in this thesis. There is, however, an ambiguity in the Philosophy for Children literature. The phrase “community of inquiry” is used in two quite different ways. It is most often used to refer to a specific classroom pedagogy, with detailed organizational and procedural guidelines, as I will outline below. In parts of the Philosophy for Children literature, on the other hand, as in Lipman’s first sentence above, the claim is made that the entire classroom can be converted into a community of inquiry. Exactly what this means is seldom spelt out in much detail. In the remainder of this section, I shall spell out and contrast the two senses, which I shall (for the moment) label the “narrow-sense” and “wide-sense” communities of inquiry, respectively. This discussion should make clear the reasons why the phrase “community of inquiry” ought to be used in the narrow sense only, and explain why it would be better to use another name for the wider sense.

Lipman (1991, 241-243) gives the definitive description of the narrow-sense community of inquiry. It has five stages: (I) the offering of the text; (II) the construction of the agenda; (III) solidifying of the community; (IV) using exercises and discussion plans; and (V) encouraging further responses. This is a very specific model of the community of inquiry, and it is the one that teachers being trained in Philosophy for Children learn. It commences with a text, written in story form, which is read around the class. This is followed by the collection of questions regarding the text from the students. The third and fourth stages both involve discussion by the class, though there is a tension between Lipman’s descriptions of each. In the notes to stage III, Lipman emphasizes the students’ responsibility for the conduct of the discussion, through the use of such phrases as “group solidarity,” “articulation of disagreements,” “joining together,” “collectively” and so on. Stage IV focuses much more on teacher control of the discussion, using phrases such as “employing questions from the academic tradition,” “opening students to other philosophical alternatives,” “focusing on specific problems” and “compelling the inquiry to examine.” Only in stage V is there any hint that the community of inquiry can take place without the students, sitting in a circle talking. Here, Lipman mentions eliciting further responses which, although they include telling, also encompass writing, painting and drawing.

Lipman’s description here captures the core meaning of “community of inquiry” in the Philosophy for Children literature. If a classroom teacher in the Philosophy for Children tradition talks of forming a community of inquiry, then the expectation would be that the children would get into a circle, read a story, gather questions and discuss

them. Of course, the narrow-sense community of inquiry can be augmented, by incorporating such techniques as pair discussion, small group work or individual writing within this general structure. Some of the exercises of stage IV encourage this, and I will allow that the term “narrow-sense community of inquiry” incorporates such enrichment, provided it fits within a methodology largely described by Lipman’s five stages. Despite this enrichment, the phrase “community of inquiry” functions very well to mark out this particular pedagogical activity from other activities that go on in a classroom, such as lecturing, writing, doing worksheets, collaborative group work, recitation discussions and so on. Together with the philosophical focus, it is what is most distinctive and, many would maintain, most powerful about the Philosophy for Children approach.

Yet we have seen that Lipman calls for us to convert the classroom into a wide-sense community of inquiry. Splitter and Sharp (1995), in their introduction, take their task to be to “explain how the classroom can be transformed into a community of inquiry based on dialogue, trust and respect.” Are we to take this to mean that students should sit in a circle all day long, reading, questioning and discussing? Is communal discussion to be the only learning technique used in the classroom? Such a recommendation would seem to be disastrous, for it would impoverish the variety of the classroom and neglect many important pedagogical techniques, as I discuss further below.

Let’s turn again to the origins of the phrase. For Peirce, the community of inquiry was archetypically an international community of scientists. This community would seldom fulfil anything like the conditions specified by Lipman for a narrow-sense community of inquiry. If scientists have texts to read, they are not stories. Rather, the texts are the world itself and the academic writings of other scientists. Nor do they commonly sit around in circles and discuss, though this does happen occasionally. Rather, the scientific community of inquiry is carried out over widely separated sites, through a combination of many, many activities, only some of which directly involve communication between scientists. Scientist who form part of the community can, amongst many other things, engage in solitary experimentation or musings, work in small research teams, give or listen to lectures, or consult experts. Yet underlying this is a crucial element of the scientific method. As Popper (1962, 220) puts it:

What we call ‘scientific objectivity’ is not a product of the individual scientist’s impartiality, but a product of the social or public character of scientific method; and the individual scientist’s impartiality is, so far as it exists, not the source but rather the result of this socially or institutionally organized objectivity of science.

It is not merely objectivity-as-intersubjectivity that is built by the international scientific community of inquiry: other products include methodologies (e.g. gene splicing), conventional standards (e.g. the cutoff points for statistical significance), conceptual schema (e.g. sea floor spreading), interpretations of mathematical formalisms (e.g. the Copenhagen interpretation of quantum mechanics) and so on. Each of these requires open communication, in which any other scientist can potentially participate, raising questions, suggesting alternatives, pointing out flaws in data, methods or analysis, supplying reasons, identifying assumptions etc., which takes place informally, at conferences or meetings, or in journals. It is in this part of the scientific community of inquiry that the parallels to the discussion stages of the classroom narrow-sense community of inquiry are the strongest. Nevertheless, while the parallels are stronger, the two communities are by no means identical. The scientific community of inquiry is methodologically much more diverse than the classroom narrow-sense community of inquiry, even if we do not go quite as far as to agree with Feyerabend's slogan "anything goes" (1975, 14).

Splitter and Sharp (1995, 24-25) discuss how "every classroom, in every discipline and subject area, can be transformed into a community of inquiry." In their discussion, they do not make specific reference to the narrow-sense community of inquiry. Rather, they talk of "redefining teaching and learning as inquiry based activities." Teachers "must tap into the conceptual bases of their respective disciplines and allow their students full scope to explore contestable issues as they arise." The lack of reference to teaching methodologies in this description leaves the methods open. Nevertheless, the narrow-sense method is undoubtedly uppermost in their minds, as is born out in the later section *Transforming the classroom into a community of inquiry* (140-147), which makes only passing reference to other methods: "small group work and other forms of co-operative learning" can be introduced into a mature community, and "virtually all forms of learning... will be enhanced" (ibid., 145) once students have internalized the procedures of inquiry.

This does not really grapple with the real problem with the concept of turning a classroom into a community of inquiry. The slogan surely cannot imply that narrow-sense communities of inquiry ought to be conducted in every session throughout the week, for this would have several bad consequences. First and foremost, students would soon become restless. Good classroom planning takes note of the students' needs for variety in classroom organization, learning styles and activities. While it is often true that children can engage in philosophical discussion for much longer periods than their teachers would anticipate, to ask them to do it all day, every day is asking for trouble. Secondly, though schools undoubtedly undervalue talk and the

construction of meaning, there are many other valuable learning methods that do not rely so heavily on pure dialogue, and cannot be efficiently achieved through the narrow-sense community of inquiry type of dialogue. Many practical skills must be learned by doing. Valuable abilities such as the instant recall of number facts are best instilled by drilling, perhaps even chanting. Large amounts of relatively uncomplicated information can often be best learned through listening to a lecture or reading a reference book. As Osborne (1996) notes, there is a role for didacticism, telling, showing and presenting.

Yet this is not to say that all these methods would not be enhanced by being judiciously interleaved with narrow-sense communities of inquiry, which could introduce or reflect upon other activities. Indeed, I have argued (§8.23) that the community of inquiry is an excellent place to build the tools of inquiry that are foundational for other learning styles. As well, the planning of such other activities, in a class which has already internalized many of the tools of inquiry, could probably be usefully carried out within a class-wide dialogue. This picture is probably closer to what Lipman, and Splitter and Sharp have in mind in calling for the classroom to be converted into a wide-sense community of inquiry, and it mirrors more closely the scientific community of inquiry.

However, even within such a classroom, as I discussed in §4.14, the teacher has a professional duty to organize the classroom, through the use of pedagogical action so that learning takes place. Here the teacher, drawing on the ideal speech situation for its regulative power (§4.124), encourages the students to enter explicit critical discourse (§4.123) in the form of philosophical inquiry (§7.3). This asymmetry in the positions of the teacher and students does not disappear (even though it might be lessened) when the class is capable of conducting a mature narrow-sense community of inquiry. At this stage, the students will have developed their communicative autonomy to a high level (§4.22). There does not seem to be any parallel to this pedagogic role in the scientific community of inquiry.

A major problem with talk about transforming the classroom into a wide-sense community of inquiry in this way is that many teachers will claim already to have done so, even if they do not utilize a narrow-sense community of inquiry at all. And it is not *prima facie* evident that they are wrong in so claiming. There are, as the Peircean scientific community of inquiry shows us, many ways of engaging in inquiry. The Philosophy for Children movement claims to have developed an exceptionally powerful and effective way to do so, but we can hardly claim to have a monopoly on inquiry. While I claim that the use of the narrow-sense community of inquiry leads to

further important and essential enhancements, teachers have created classrooms in which the spirit of inquiry is strong using many different specific teaching methods.

The conflation of the idea of the (narrow-sense) community of inquiry as a specific method for fostering philosophical discussion and critical discourse, with a (wide-sense) classroom-wide ideal for the transformation of education, is a confusing and unnecessary one. What is distinctive and important about the narrow-sense community of inquiry risks being lost. As I have said, a narrow-sense community of inquiry can be usefully enriched by incorporating other methodologies, such as using materials other than stories as the trigger, breaking into small groups or pairs for specific tasks, or adjourning so that individual research can be carried out, before reconvening to discuss the outcomes. Indeed, the use of exercises and activities in Philosophy for Children support materials (see stage IV in Lipman's characterization of the community of inquiry above) already encourages such enrichment. But I propose henceforth to restrict my usage of "community of inquiry" to those classroom activities that largely fit with such enriched versions of Lipman's narrow-sense description. The concept of a classroom converted into a wide-sense community of inquiry ought to be marked by a different phrase, and I propose "the inquiring classroom community."

In the remainder of this section, I want to expand and defend the contention that the community of inquiry, in the sense just delineated, ought to be used in classrooms as a core element of a program of moral education. In this, I shall draw further on the philosophical positions developed in Parts I and II.

9.1 The Development of Reasonableness

In §3.12, I identified five aspects of reasonableness: the critical, creative, committed, contextual and embodied aspects. To be reasonable, I argued, is to be able to incorporate all five aspects into one's thinking in a way which balances the imperatives of each, and does not give too great an emphasis to any one aspect. Due to this multi-aspectual nature of reasonableness, we are unable to assert that there is one unique and decisive way to think about many complex problems (§3.126). Thus the idea of a Unitary, Universal Reason is not merely practically unattainable, but also ideally incoherent (see also §6.232).

Reasonableness, however, must be one of our educational targets. It encompasses the intellectual virtues (§7.25) and is a corequisite for the possibilities both of becoming reflectively morally virtuous (§5.2) and becoming more autonomous (§4.2). As we have seen, reasonableness is multi-aspectual (§3.12) and does not fall prey to the problems that beset Unitary Reason (§6.2322).

Our pedagogical methods need to recognize this. Any method for developing reasonableness ought to invoke all five aspects as appropriate. Yet many thinking skills courses focus largely on critical reasoning, often breaking it down into many isolated skills which, it is thought, can be assembled into complex thinking. Hence, these sorts of programs teach formal logic, informal fallacies, identification of assumptions, argument structures, and so on through demonstration, worked examples and repetitive practice. Such teaching is commonly done on toy examples, specifically constructed to fit the skill being taught. These examples are usually quite divorced from the lived lives of the students, and almost never chosen by them.³

In a review of 27 studies on the influences on college students' critical thinking development, McMillan (1987) found that very few showed any significant changes due to specific courses or instructional methods, although they did conclude that weak research design in most cases meant that the conclusion that critical thinking courses were ineffective might not be warranted. There was, however, good evidence that college attendance as such did improve critical thinking, though the factors responsible had not been identified in these studies. He quotes Brabeck's (1983) opinion that critical thinking skills are necessary, but not sufficient, for the attainment of high levels of reflective judgement. In his discussion, McMillan claims that the evidence suggests that good critical thinking courses would have an everyday emphasis, a meta-cognitive focus and be set in specific content domains.

Pascarella and Terenzini (1991, Chapter 4) report similar conclusions about the effectiveness of critical thinking programs, and about general college attendance to McMillan's. Factors they identified as associated with gains in critical thinking fell into two categories: general factors to do with activities at the college; and factors associated with classroom organization in specific courses. In the former category, they include following a general education curriculum, being socially and intellectually involved in college life, and having informal, non-classroom contact with lecturers. Within classrooms, the degree to which teachers encourage, praise or use student ideas; the degree to which students participate in the class; the extent of peer-to-peer interaction; an emphasis on high cognitive level student discussion, problem-solving or learning-cycle/inquiry techniques; and small class size were all associated with improved thinking.

From studies such as these, we can gain some clues as to which factors to look for when trying to find effective methods of improving the thinking of students. McMillan's advice that thinking needs to be situated both in everyday settings, and

³ This characterization of thinking skills courses owes much to van Gelder (1998).

within specific content domains, is supported by my analysis of reasonableness, for these factors connect to the committed, embodied and contextual aspects of thinking. If thinking is richly contextualized in a way that connects to the everyday lifeworld of the students, then this will entail that the committed and embodied aspects of thinking will be enlisted. If the context creates real puzzles, and the students are motivated to try to solve them, then both the critical and creative aspects will also be involved. If the means of dealing with these is relatively open-ended and student-driven, then the aspects will be invoked when the discursive context requires them, and they will be incorporated in an integrated way. The classroom list of Pascarella and Terenzini (with the possible exception of the small class size) reads like a list of tips to a teacher in a community of inquiry.

The community of inquiry approach used in the Philosophy for Children program matches many of the features identified above. Starting by reading a story taps in to the strengths of narrative, as a way of grounding the discussion in a context recognizable to the participants. Since the students set the agenda for the discussion by asking their own questions, the link to their own interests and lifeworld is strengthened. Thus they will engage not just critically, but also emotionally and viscerally, making connections through the offering and analysis of anecdotes (this will be further developed in §10.223). Student ideas and speculations form the grist of the discourse, and are analysed, modified, accepted or rejected, on the basis of a rich investigation of the reasons for and against them. Conceptual analysis encourages students to explore imaginatively the concepts involved, so learning how they can be metaphorically extended in concrete situations (§7.23). Student participation, particularly peer-to-peer, is paramount, though, as we saw in §4.141, the teacher's pedagogic action entails strategic action that ensures the discourse attains sufficient rigour. This scaffolding is an essential element of Vygotskian learning.

The case presented here supports the thesis that a classroom community of inquiry provides an effective means for developing reasonableness in all its aspects. The claim that Philosophy for Children has positive impacts on children's thinking has been supported in a series of empirical studies, although the focus of these studies has been almost entirely on the critical aspect of thinking, with some minor attention to the creative and committed aspects. Fourteen are summarized in Lipman (1986), and another four in Gazzard (1988). Between them, these studies found that classes using Philosophy for Children materials showed gains statistically significantly greater than control classes in reasoning ability (including formal inferences, finding alternatives and possibilities, providing reasons and ideational productivity), reading comprehension, reading and maths achievement, moral justifications, self-concept and

teacher perceptions of their ability to function rationally. In my own work (Sprod 1994c; 1997c, d; 1998), I was able to demonstrate that a Year 7 science class that incorporated a Philosophy for Children approach showed significantly greater gains in scientific reasoning when compared to a control class, and that the Vygotskian theory outlined above best accounted for the gains (see §8.21). I will return to this study in §10.222, where I present some empirical evidence that all the aspects of reasonableness were used by students in those discussions.

This is true independent of the specifics of its content matter, provided it is rich enough for inquiry. As Splitter (1997) argues, the investigative nature of the community of inquiry ensures that, “irrespective of whether the inquiry engaged in by a community is philosophical, the community itself is philosophical.” He makes this claim on the basis that any community of inquiry at any depth will soon have to deal with common, central contestable concepts: common in that they are found in everyday talk (fair, good, beauty etc.); central in that they are vital to the way we make sense of our experience; and contestable in that they “resist our best efforts to define them with clarity and finality.” Such concepts “help to establish philosophy as a discipline with its own content or subject matter,” so that a philosophical investigation is not just a process, but a process that includes engagement with this sort of concept.

The procedure, of course, is as important as the concepts, for to deal with the concepts in a manner that is not reasonable (in my sense) would be not to engage philosophically with them. Reasoning is at once the tool-kit and part of the content of philosophy, for a reasonable inquiry into concepts will not erect barriers that prevent investigation of the *method* of the inquiry. Inquiry will turn reflexively onto its own methods, and thus introduce the meta-cognitive element that McMillan (1987) identified as being so important. Yet this reflexive investigation of the tools of reasonableness sits firmly within a context, and since the tools are being honed to serve a present need in an inquiry chosen by the participants, it is their contextualized usefulness that is established. Thus is the critical aspect of thinking dependent on the contextual aspect. A community of inquiry will use and reflect on the same tools many times in different contexts, thus encouraging the sort of generalization from context that leads to more domain-general reasoning.

Thus, both philosophical and empirical evidence supports the contention that a community of inquiry is a powerful pedagogical method for fostering reasonableness, though future empirical work ought to pay more attention to all the aspects. Indeed, since the community of inquiry is in a sense merely a formalization and intensification of Habermasian communicative action and critical discourse (§4.123 and §4.124) that

form reasonableness in the lifeworld, deliberately scaffolded by the teacher through pedagogical action (§4.141), there is every reason to believe that this would be so.

9.2 Ethical Development in the Community of Inquiry

In the preceding subsection, I made the case for the community of inquiry as a pedagogy for the development of reasonableness, and I asserted that the specific content of the inquiry is not critical, provided that it engages with Splitter's common, central and contestable concepts. In this subsection, I will focus more closely on communities of ethical inquiry. I will investigate the contribution that the engagement of students in such inquiries can make to their development as ethical persons.

It is important to be clear from the start that I am not making the hegemonic claim that all that is needed to produce ethical persons is to engage them in a community of ethical inquiry (in the pedagogical sense that I am using here). There are three things to say about this. Firstly, as I discussed at the commencement of Chapter Eight, neither classroom teachers nor schools as a whole have a monopoly on moral education. Secondly, even within a classroom, the teacher uses many other means of moral education, such as listing, explaining and enforcing a set of classroom rules, enforcing and reasoning about informal or implicit expectations of behaviour and so on (§8.1).

Thirdly, there is a broader sense in which we are all engaged in an informal, society wide community of ethical inquiry. This is seen every time a child asks "Why?" when required to act ethically, or when a parent offers a reason for requiring behaviour even without being challenged. It is a Habermasian engagement in communicative action, possibly transformed at times into practical discourse, that is an essential part of the lifeworld. As we have seen, this blurs any sharp distinction between moral education and moral training. This brings us closer to some of the central concerns of this thesis, for as we saw at the conclusion of §8.23, this type of moral development of the young can be equally well described in Habermasian language, or using the Aristotelian language of habituation into the virtues mediated by practical wisdom (§5.21).

This last point, however, does highlight one of the most important features of the community of inquiry as a pedagogic method. It is not a totally new and strange method of learning to the children who take part in it for the first time, because it formalizes and makes more explicit the communicative action and discourse that children will have engaged in frequently. Here, Habermas' account has the advantage over Aristotle's account which (as we saw in §5.21 and §5.23) has been interpreted to mean that children cannot take part in philosophical inquiry. For Habermas, the

young are not excluded from engagement in practical discourse. Indeed, they may enter into practical discourse (albeit at a low level of competence) whenever they, or others in their dialogical group, call a validity claim into question.

It is this familiarity that provides much of the power of the method, and it also leads to students claiming that they are not doing any work while they are participating in a community of inquiry. Nevertheless, there are important differences as well. The formalization, the larger group than is normal in lifeworld conversations (which can lead to problems with the dynamics of turn taking), the more explicit focus on inquiry and the fuller coverage of ideas expected are all features that can make a community of inquiry strange to students. As the discussion in §4.141 made clear, the major difference lies in the pedagogic action of the teacher. Communicative actions and discourses within the lifeworld do not generally contain participants who are moulding the conversation for concealed ends, for if they did, Habermas would claim they were examples of strategic action instead. Yet when a more experienced person - teacher, parent, coach etc. - is involved with less experienced persons in an instructional role, they must through pedagogic action include some strategic actions (even if the ends need not remain concealed once they are made a subject of the inquiry). It is part of their pedagogic role, and this is especially important in a community of inquiry.

A community of ethical inquiry takes ethical concerns as its content. This means that the set of concepts to be raised by the initial story will include a large number of the common, central and contestable concepts of ethical life. We can expect stories in which the characters find themselves in situations with ethical implications, wondering what is the fair, right, good or acceptable thing to do. We can expect them to be considering their own actions and reactions, reflecting on what sort of person they wish to be. For this, they will simultaneously need to develop their reasonableness, for this is essential to the ability to engage in such considerations. The situations in which these concerns arise will not be extraordinary ones, but rather they will be set in the sorts of narratives that might engage ordinary children in their everyday lives. Yet the outcomes of the story will not be clear cut and unproblematic: these are not fables with neat morals at the end, like Aesop's fables. Nor will the philosophical puzzles that arise be restricted to ethical ones, for life is not so neatly compartmentalized either.

The Philosophy for Children canon contains quite a number of stories of this sort. First (chronologically) and foremost (in terms of length and depth) among them is Matthew Lipman's novel *Lisa* (1983), with its accompanying manual *Ethical Inquiry* (Lipman and Sharp 1985). It is worth looking in some detail at what Lipman has to say about ethical inquiry in his Introduction.

Lipman characterizes ethical inquiry as an “objective and dispassionate inquiry into moral problems and moral situations. Its aim is never to indoctrinate, but rather to help people more clearly understand what their moral options are and how those options can be critically assessed.” He contrasts ethical inquiry with values clarification, decision-making and moral-dilemma stage theory, averring that each has something to offer moral education, but none covers the whole of it. To identify moral education with decision-making or moral-dilemma stage theory is, he avers, like identifying farming with harvesting. Although harvesting is essential in farming, it is by no means the whole of it. For good moral decisions to be made, the ground must be prepared first, and so Lipman claims that

a sound moral education minimally involves helping children understand: what criteria are and how they function; the significance of assumptions; the process of reasoning; the giving of good reasons; the moral character of situations; the relative importance of and proportion between parts and wholes; the opinions of other people; the interests of the community in which one finds oneself; the need to take all relevant factors into account; the need to weigh consequences; the importance of neither overestimating or underestimating the role of the self in the context of a moral situation; the importance of sizing up other people's and one's own intentions; the anticipation of possible harm as a result of one's actions, both to others and to oneself and[;] the fundamental importance of preventing moral crises before they occur.

Such a project is, according to Lipman, long term and open-ended, preparing students both to adapt to change, and to hold on to the valuable in a changing and pluralistic world. He asserts that:

The aim of ethical inquiry is not to teach children certain particular values; it is rather an open-ended, sustained consideration of the values, standards and practices by which we live, discussed openly and publicly so as to take all points of view and all facts into account. It is the assumption of ethical inquiry that such discussion and reflection, taking place in an atmosphere of mutual trust, confidence, and impartiality, can do more to foster moral responsibility and moral intelligence in children than any system which merely acquaints them with “the rules” and then insists that they “do their duty.”

This account fits well with the case developed in this thesis, though I shall comment on one discrepancy shortly. Lipman, like me, sees ethical education as a wide ranging and long term endeavour. While his list of necessary conceptual capacities shows that an account of rationality essentially underlies ethical inquiry, this list is expansive, and includes important creative and contextual elements. The rationale for such an emphasis is particularly acute given the diversity of modern society. With his farming metaphor, Lipman follows Aristotle in claiming that ethical action springs from a reflectively modified character immersed in context. Like Aristotle and Habermas, Lipman recognizes that we all live in a value-laden lifeworld, habituated to certain values, standards and practices, and that the purpose of ethical education is to strengthen the ability to reflexively assess them. His emphasis on open and public

discussion echoes Habermas' insistence that this reflection must be dialogical, not monological and, though he doesn't explicitly say so here, in *Thinking in Education* (1991, Chapter 15) he makes it clear that this will also prepare the child to enter the larger discourses of democratic society. The reference to an atmosphere of mutual trust, confidence and impartiality shows that ethical inquiry must be grounded in respect.

The one detail with which I would disagree is in the description of the inquiry as "objective and dispassionate." Lipman is concerned that the inquiry ought to be aimed at reaching beyond the idiosyncratic views of any one individual, and that the moral positions advanced ought to be supported reasonably, and not through displays of partisan emotion. However, in the light of my discussions of agreement (§6.21) and reasonableness (especially in §3.123), a better description of the inquiry would be that it ought to be tolerant and appropriately passionate. The committed aspect of thinking is essential if judgement is to be linked to action and this requires passion, but inappropriate displays of intolerant passion will not assist the inquiry. Objectivity can only be understood in the sense of the search for intersubjective agreement, either in consensus, or in agreement to disagree.

As we have seen, following Aristotle, to be ethically good is to act in accordance with virtue, and this requires doing the right thing in the right way at the right time (§5.2). Deciding the right thing is something that arises from character. So, if we are to be responsible for our actions, we must be responsible for our characters. If our character were to be merely the product of our upbringing, then we could not be responsible. Yet in the absence of obvious ethical disability, our society will hold us responsible. Now, my account of habituation (§5.23) shows that there is no such thing as a mindless habituation, for even habitual actions need judgement concerning context and suitability. Yet, even if we have been habituated well, it is possible to be relatively unreflective in our ethical actions and hence to be prone to make poor judgements.

One essential element of moral character, then, is the ability to read ethical contexts well and to act reflectively according to virtue. In frequently occurring ethical contexts, these responses can be automatized, but to be a morally mature person requires the ability to withdraw from action in order to reflect when circumstances are such as to call these automatic responses into question. This reflective ability itself, as I have argued, is parasitic on engagement in communicative action and (especially) critical discourse. Although we all do so engage in our ordinary lives, schools can provide opportunities both for building the capacities of reasonableness on which reflection is based, and for deliberating collectively about the ethical import of many

different factors in different contexts. The community of ethical inquiry is an ideal setting for both.

Collaborative reflection is important for morality not only because our reasoning is improved through Vygotskian learning, but also because ethical decisions almost always concern our relationship with others. Our ability to frame a situation morally depends crucially, as Lipman notes, on our ability to read others in various ways and incorporate consideration of their interests. This too cannot be learned monologically. Exhortations to “do as you would be done by” are worthless (or even harmful) unless we are able to imaginatively situate ourselves as we would be situated if we *were* the other, rather than if we were ourselves in the other’s place. Developing Arendt’s “visiting imagination” (Glaser 1998) is something that can only be done dialogically, and the community of inquiry, with its dynamic aimed at understanding the other, enables Vygotskian learning of this capacity too.

Schools are places where disembedded (or decontextualized) thought is both valued and taught (Donaldson 1978, 76-78). For this reason, perennial complaints arise that school knowledge is not transferable to the “real world” (see also Claxton 1991; Solomon 1992, for discussion in the context of science education; Splitter and Sharp 1995, 70-79). Similar difficulties arise in moral education, where general principles of morality can be taught, but remain somewhat disconnected from real world situations. This is, of course, exactly the problem identified by Habermas when he makes the distinction between discourses of justification and discourses of application, as discussed in §7.24.

While we might all agree that it is wrong to tell a lie, and that we ought always to be honest, just how to apply this in specific situations is underdetermined by this agreement. Johnson (1993, 91-98) discusses at length the reasons why this should be so, given the prototypical structure of the concept “lie.” Central examples (the deliberate attempt to mislead the hearer using a statement the speaker knows to be false, in such a way as to disadvantage the hearer and advantage the speaker) are relatively unproblematic, but examples that are somewhat like this in some respects, but unlike in others (a white lie to protect the hearer’s interests, for example) are more difficult to deal with. Pritchard (1996) also discusses the case of lying and honesty at length, comparing the simplistic and moralistic cautionary tale told of George Washington, who allegedly claimed as a child “I cannot tell a lie, father,” with the picture book *Honest Andrew* (Skurzynski 1980), in which young Andrew Otter comes face-to-face with the difficulties of trying to live up to his father’s advice that “an otter ought to be honest.”

As Pritchard claims, it is by the dialogical exploration of stories like *Honest Andrew* in a community of inquiry, rather than by blindly following George Washington's simple rule, that children can come to grips with the tension between abstract, decontextualized moral principles and situated moral action. Pritchard invokes the notion of casuistry, or case-by-case reasoning. If we turn to the exploration of morally puzzling cases with children, he claims, then moral development is enhanced by the search for agreement, the development of shared principles, which are nevertheless open to revision from new cases, the sorting out of relevant similarities and differences, and by attention to nuances. Such development needs good character as well as good reasoning, and open discussion to build familiarity with context. In Habermasian terms, a community of inquiry must by turns enter discourses of justification, to generate moral principles (i.e. rules of thumb for moral judgement) and discourses of application, both to hone the ability to invoke them in the context of lifeworld applications, and to test these principles against problematic cases. A student who can do this is communicatively autonomous. In Aristotelian terms, virtue cannot be identified with the following of mechanically habituated virtues. To be virtuous requires situated practical wisdom together with caring about the sort of person one is (§5.2). If one does not approve of one's habits, one has a responsibility to habituate new virtues. Thus, under a virtue ethics account, do children become morally autonomous. The Habermasian story, however, makes it clear that such autonomy is parasitic on the intersubjective development of the requisite situated rationality (i.e. reasonableness).

In the pursuit of well-being under virtue ethics, rightish questions of how to treat others justly and goodish ones of how to serve my own best interests intertwine (see §6.24 on the rightish/goodish distinction). Children, and particularly adolescents, are very interested in the intersection of the two. As Lipman and Sharp (1985) point out, they are not only concerned with what is right in total, but also with what is right *for me*. Many adolescents are torn between a strong moral sense and a burgeoning individuality, trying to discover or invent who they are. In the light of this, questions of whether a society has a right to impose standards on its members, and whether individuals have a right to flout them, are live and contentious. No easy line can be drawn between the right and the good, and the community of inquiry is a safe haven in which to negotiate these burning issues in their lifeworld. Moral maturity and autonomy requires the ability to integrate the two (see Gilligan 1982, Chapter 6, for an interesting parallel account of maturity as the ability to integrate separation and connection).

I have made a case that the community of ethical inquiry ought to be an essential part of a program of school-based moral education, on the broad grounds that it improves children's abilities to be reasonable about ethical matters. Yet there are two possible objections to a reasoning based approach to moral education. Firstly, if one decides in discourse that certain things ought, or ought not, to be done, this does not entail that one will actually do, or avoid doing, those things. Secondly, practice in rationally deciding that certain actions are right or wrong may merely lead to the ability to rationalize one's actions as right, whether they are or not. I characterize these problems as the problem of turning judgements into actions, and of bush lawyerism, respectively, and will discuss them below.

9.21 Judgement to Action

Socrates famously held that if you know the right, you will do the right, to which Aristotle replied that this is "manifestly at odds with the observed facts" (see Rowe's discussion, 1993, 125-126). Certainly, we all have experience, even if only minor, of knowing what we think we ought to do, and of nevertheless doing what we think we ought not. There are two possible reasons for this outcome: weakness of the will; and knowing the good but choosing to do evil. I do not wish to explore this problem here, but merely to point out that it is a perennial problem for all moral agents, and that it is unlikely to be solved by a particular form of moral education.

Nevertheless, it is a reasonable question to ask of any proposed moral education program whether it contains the resources to strengthen the link between judgement and action. A program which turned out persons who could unerringly identify the right thing to do, but who seldom actually did it would not be producing morally good persons. In relation to the community of ethical inquiry, I shall offer two sorts of answer to this question. Firstly, I have repeatedly pointed out that the community of ethical inquiry cannot stand on its own as the whole of a moral education program: it must form part of a wider package. Secondly, it does have features that give it an important role to play in linking judgement to action.

The wider package to which the first point above refers includes elements that recognize that the carrying through of judgement into action requires practice. Consequently, a total school-based moral education package ought to have elements that are directly related to actions with moral import. This will include a social setting in which students are expected to take responsibility for their actions, and which imposes rewards and sanctions for actions which are respectively morally commendable or morally blameworthy. It will also include opportunities and encouragements to perform morally desirable actions. Additionally, teachers and

others in positions of authority within the school will provide good role models for ethically good action. This admittedly broad brush sketch of the other components of a full moral education program is not meant to be exhaustive or detailed, as my focus here is on the community of ethical inquiry. I merely mean to indicate that other elements are also essential.

I cannot be so sketchy with respect to the other answer. The resources the community of inquiry has for strengthening the link between judgement and action are important to my argument that it ought to be used in a moral education program. Before doing so, it is worth noting that much moral judgement is not explicit. In relatively unproblematic situations (or situations we perceive as unproblematic), moral actions are largely habituated or automatized. This is the case in many everyday moral situations, and if we have built a good character, we make tacit judgements about the right thing to do, and then do it. The question of linking judgement to action thus has two prongs: What do we do when we are faced with a morally charged situation that is sufficiently unusual that we are forced to make an explicit rather than tacit judgement? How do we react when we find that our tacit judgements (i.e., our habituated reactions) are not leading us to take actions about which we are sure we approve?

The first, and possibly narrowest, way in which participation in a community of ethical inquiry helps to link judgement to action becomes evident when we note that students in such a community are acting every time they contribute to the community. Speech is an act in itself, as are such forms of non-verbal communication as nodding in agreement, looking disdainful and so on. In a community based on respect, actions and utterances which are respectful of others are encouraged and reinforced. In the initial stages of building a community, these respectful actions may be continually encouraged by the teacher, or made the subject of negotiation within the community. Once the community is well established, they will be largely habituated. Nevertheless, students will be able to make respectful or disrespectful actions within the community once more the subject of the discussion, and hence be able to evaluate the reasoning that leads to the actions.

It is important to realize that this account is restricted to action and judgement within the community itself. Since we know that competency is contextualized, we cannot expect that students who can link moral judgement to moral action excellently in the classroom community of ethical inquiry will unproblematically transfer this competency to their everyday lives. It is for this reason that the community ought to form part of a wider moral education package, as discussed above. Nevertheless, there are still other resources within the community of ethical inquiry that can assist this task.

Damon (1996), pointing out that empirical studies find that it is rarely the case that moral views are completely melded with an image of one's self, says that this fact implies that moral judgement not in itself a good indicator of morality. Morally praiseworthy persons, he claims, need both moral beliefs and also a strong personal identification with them, so that they have a place of great importance in one's sense of self. The community of inquiry often focuses on exploration of what it is to be a person (Splitter and Sharp 1995, 167-171), and thus, through reflexive consideration, each participant will reflect about the sort of person they are becoming: a view that parallels Aristotle's (§5.2), but gives it (as does Habermas) a dialogical twist. The habituation for linking thought and action practiced within the community can be carried over into the wider community through seeing one's self as a character in an ongoing story.

Philosophy for Children emphasizes narratives, and narratives, as continuing stories, frequently link the judgements that characters make to their subsequent actions. There are two levels on which this can happen: firstly, for the characters in the stories used as trigger materials for discussion and, secondly, for the students of the community itself, in that they include anecdotes about their own lives into the discussion.⁴ Indeed, when participants inject into the discussion a story from their own life, they are not only able to reflect more explicitly about their own actions, but they also provide to the other participants the opportunity to see into others' lives. In Vygotskian fashion, this strengthens their moral imagination. Jointly, the community can explore an issue through taking into account not just disembodied reasons for abstract cases, but multiple perspectives on situations drawn from the lifeworld of their peers. This continuing symbiosis between conceptual exploration and lived experience aids transfer from discussion to action.

The linking of the fictional narratives to lived narratives is more likely to happen the richer the fictional narratives are. We can consider Kohlberg's moral dilemmas as narratives, but they are thin, devoid of interesting contextualization about the characters and situations. For this reason, some of those confronted with them (women more often than men), find it very difficult to come to judgement in the absence of such information (Gilligan 1982, 100-101). The longer stories and novels used in the Philosophy for Children program allow for the building of context and deeper understanding of the characters of the fictional characters in the story (indeed, it is hardly coincidental that the word "character" is used in both these senses).

⁴ The place of anecdotes in a community of inquiry is further explored in §10.223.

The richness of the narratives used in the community of ethical inquiry is reinforced by the rich account of reasonableness that underlies the community. Instead of a narrow concentration on the formal correctness of arguments that lead to judgement, having a multi-aspectual understanding of reasonableness leads to a stress on integrating critical, creative and committed thinking in the light of specific contexts and visceral reactions. Certainly, there is a need to be able to monitor arguments for validity, but without the capacity to imaginatively extend our prototypical understanding, or to put ourselves in others' shoes, we will not be able to find ways of morally considering the most interesting and puzzling morally charged situations (see Johnson 1993, Chapter 8, for a full account of the place of moral imagination in moral thought). If we rigidly exclude emotions from our moral thought, then, as Hume so rightly pointed out (see §3.123), there is no way to motivate even the most logically impeccable moral judgements (or even the desire to come to judgement itself).

Nor should we seek to separate out the considerations of a community of inquiry into episodes that consider each of these aspects independently, although there might be times when one or other is given greater emphasis. Certainly, we will sometimes want to abstract to a degree away from particular contexts, as we enter discourse concerned more with justification, but we will need to take much more notice of contextual matters as the discourse veers towards application. Sometimes, we will need to concentrate fully for a while on the structure of an argument, or the reasons why we feel strongly about a certain matter, but the conversation as a whole will need to cover these and other matters as the discussion requires.

Without the capacity to reflect on their own actions, especially in a way that incorporates a well grounded ability to envision the ways in which the particular others who are involved might see them, children will not be able to evaluate their actions well. Communal reflection is the method by which such individual reflection can be best built (§3.23). This ability to reflect is certainly not a sufficient condition for moral action, and neither, at least in simple, prototypical cases, is it necessary. But it is necessary for taking the sort of care about our character that Aristotle identifies as being so important (see §5.2). In other words, it is important if we are, as moral agents, to approach autonomy (§4.22).

9.22 Bush Lawyers

A "bush lawyer," in Australian usage as defined by the Macquarie Online Dictionary, is "one who attempts complicated and often specious arguments to prove a point." The worry here is that if we improve students' abilities to reason about morality, they will use this ability to excuse their actions even when those actions are not morally

good, and they know they are not. This is often expressed by saying that they are offering, not reasons, but rationalizations for their actions.

Let's try to be clearer about the distinction between reasons and rationalizations. The key question is whether the person offering the reasons really believes them. If they are offering reasons duplicitously, then these are not reasons, but mere rationalizations. It can also be the case, however, that reasons are employed to bolster a certain degree of self-deception. If the self-deception is, at a more basic level, being knowingly practiced, these reasons are rationalizations.

An example will make these points clearer. Fred is found shoplifting a small item from a bin on the pavement outside a shop (I shall assume for the sake of this example that this is in fact an immoral act, and that Fred knows this). When asked why, Fred offers a number of reasons: the shop is part of a big chain who can afford it; the shop rips people off anyway, so we ought to rip it off in return; if they place temptation in people's way, they are asking for things to be taken; capitalism is evil, so stealing from capitalists is morally acceptable. If Fred knows that shoplifting is clearly a case of stealing and thus morally wrong, then his "reasons" are really rationalizations. Fred is just trying to talk his way out of trouble. Alternatively, Fred might really want the item, and be looking for some way to convince himself that it is really all right to take it. If he heard someone advance the above reasons and pretty well convinced himself that he believed them, then he would be self-deceptively rationalizing his act. Finally, Fred may be a member of a group who genuinely believes these arguments. For him, these are good reasons and hence not rationalizations, even though, for us, they are bad reasons.

If, through the community of ethical inquiry, we improve a child's ability to reason in moral matters, this reasoning can be used genuinely, duplicitously or self-deceptively. Used genuinely, it will improve that child's ability to identify the right and wrong acts. But we cannot rule out the possibility that it will be used duplicitously or self-deceptively, nor do I think that there is anything we can do to guarantee that this will not happen. The question is: is there anything in the nature of the community of ethical inquiry that will tend to decrease the likelihood of either of these things happening? I believe that there is.

The key to this assertion is the *community* part of the community of inquiry. Participants within the community have responsibilities to the community. These include the responsibility to try to establish their position if it is challenged, the responsibility to challenge another's position if it seems problematic, and the responsibility of respect for each other. In an established community, these

responsibilities will be taken seriously, and in a new one, the teacher will insist that they are.

The upshot of this responsibility is that it ought not to be easy to get away with duplicitous reasoning (provided, of course, that the whole community is not similarly duplicitous on this point). Any student who believes that the position put forward by another is wrong can, and has a responsibility to, challenge that position (or, failing a student challenge, the teacher must wield pedagogic action: see §10.23 for further discussion). In the ensuing discussion, rationalizations will have to compete with reasons. While there is no guarantee that the reasons will expose the rationalizations, the challenge will at least ensure that the duplicitous reasoner will either have to construct a strong and coherent case through rationalizations, or admit that the case will not stand. In this need for public justification, the *bona fides* of the participants can be tested, and the duplicitous player will have pressure brought to bear to return to genuinely held views.

Self-deceptive views can be genuinely held, of course. The participant who offers reasons in defence of a view that is based on self-deception will not be deliberately trying to “pull the wool over the eyes” of the other participants. Nevertheless, the challenge that their views are likely to provoke will lead to a closer look at some of the assumptions that underlie the view advanced. In this deeper examination of views, and of the person that propounds it, greater self-transparency will be gradually constructed. The self-deception will be replaced with a clearer self-understanding. Of course, the actor might now turn to other-deception and hold onto the view in this way, but in most cases, the disappearance of self-deception will lead in the other direction.

My discussion so far has applied to the classroom community only. The broader society into which the students go is not a community of inquiry, and none of the responsibilities I have outlined might hold. It seems that it might be possible for a student to learn reasoning within the community, and be constrained there to avoid duplicity, but to use specious reasoning quite deliberately when others do not question. Clearly, if the education system as a whole equips students with Postman and Weingartner’s (1969) “crap detector,” then the conditions for practicing duplicity in the wider community would be lessened.

There is an additional safeguard in the method of the community of inquiry, for the charge that teaching moral reasoning only produces bush lawyers assumes that good reasoning in this argumentative sense (i.e. largely the critical aspect) will be learned independently of any moral flavour. In a community that works on the rich model of

reasonableness I have outlined, discussions will be shot through with the imaginative empathizing with others that lies at the core of moral sentiment, and investigations of one's own, and others', emotional commitment that drives good thinking. If such thinking is built publicly by the community, then the Vygotskian internalization of this thinking will ensure that these aspects are well integrated in the individual thinking of the students. By contrast, thinking that is taught in a disembodied, decontextualized way will make this sort of separation of good (as in efficient) thinking from good (as in morally acceptable) thinking much more likely.

My account is a little too simplistic. Not all (or even most) disagreement within the community implies that somebody is being duplicitous or deluded. The respect which must be taken as a foundational ideal within the community (§6.24) means that neither side ought to level one of these charges at the other without good reason. Caring for another does not always mean shielding them from attack, even though it might often do so, but it does influence the manner in which such an attack might be made.

In summary, any time we teach children to reason better, we run the risk that they may misuse that reasoning. The best safeguard against this outcome is to teach reasoning in a way that avoids divorcing it from its moral uses. The community of ethical inquiry meets this condition. For if we want to educate children so as to become moral persons, we must teach them to reason. A person with poor reasoning may be crudely habituated to "do the right thing" in some centrally common situations, but they will lack the capacity to be a fully functioning moral agent with a large degree of moral autonomy in situations with any degree of moral puzzlement or import about them (§5.21).

9.3 The Place of the Teacher

So far my account of the community of ethical inquiry in the classroom has concentrated largely on the students. Yet, as we have seen especially in §4.141, the teacher has a central and important role in the community. The teacher must wield power via pedagogic action if the classroom is to be transformed into a community of ethical inquiry in the first place, since children left to their own devices would be most unlikely to spontaneously form such a community.

The teacher inevitably occupies a unique place in the classroom. Students do not, and probably cannot, consider the teacher as like just another student. Teachers are expected to try to organize the classroom, and to take responsibility for the overall conduct of the room. This is not to say that the teacher always has the *de facto* power in a classroom, for the students may usurp that. But the teacher is the *de jure* power, and this position means that, however teachers act, whether they have effective power

or not, their behaviour will have a special (quite possibly greater) impact on the students than the behaviour of another student.

I will not discuss the case of the teacher who has greatly diminished authority within the classroom, for it would seem to be impossible to create a community of inquiry under such conditions. Rather, I intend to analyse the place of the teacher in a properly constituted community of ethical inquiry. In such a classroom, we can assume that the teacher has the respect of the students in the class, and that respect extends also to fellow students. We cannot, however, assume that the same type of respect is extended to the other students as to the teacher. Indeed, it will differ, for the students will respect, among other things, the authority of the teacher, whereas they will not attribute the same sort of authority to each other.

In what does this authority consist? Clearly, one of the prime areas of authority recognized by the students in the teacher is organizational authority. While students have a clear right to request another student to do something, the teacher has the authority to demand it. Students can suggest classroom activities, topics of discussion, work to be done and so on, while teachers can, indeed have a duty to, take overall control of these matters. It is teachers who in the end decide on the delivery of the curriculum, even if they derive their authority from the educational system and may at times (whilst retaining a veto) devolve some of this authority to the students.

To leave the picture at this would be to imply that the teacher's authority is purely coercive (as indeed it is in part). Organizational authority that depends on such factors as age and external power structures can be seen as somewhat arbitrary. But in reality, the power of a teacher rests also on epistemic authority: that is, the teacher is not only *in* authority, but also is *an* authority. Students respect teachers and willingly cede power to them largely insofar as they respect both the knowledge that teachers have and the ability to impart that knowledge. Teachers largely have their organizational power delegated to them on the same basis.

The epistemic authority of a teacher ensures that any view that teacher expresses will (in the absence of good reasons to the contrary) be given a far greater weight than would be given to the same view if expressed by a student. That there is a real asymmetry between the epistemic authority of the teacher and the student can be seen in the fact that a teacher may almost always flatly contradict a student's assertion without anyone thinking it impertinent, whereas the opposite is certainly not so. In many cases, particularly when the teacher's view concerns matters of relatively uncontested facts, this greater epistemic weight for teacher utterances is perfectly

reasonable. However, when the view is more contestable, as is often the case in matters of value - and with many facts - the matter is more complex.

I do not wish to enter into the argument concerning the fact/value distinction, other than to note that some alleged “facts” can be very contentious, while some values can be widely accepted and uncontentious. Indeed, teachers and schools assert and reinforce certain values, such as the importance of listening to the teacher, or the unacceptability of hitting other students in class, on a regular and largely uncontested basis. These values can be quite explicit (appearing on lists of classroom rules and so on), or they can be implicit, perhaps being promoted by the teacher without explicit awareness that this is what is being done. These “attitudes and values [implicitly] conveyed by the content or the process of learning” have been “referred to as the ‘hidden curriculum’.” (Henry, Knight et al. 1988, 60). These hidden curriculum values are promoted by the school with relatively little overt discussion of their suitability.

Many of the values of the hidden curriculum are not at all contentious, even when they are made explicit. For example, the cluster of values associated with respect largely fall into this category, although there may be many contentions about how exactly respect ought to be manifested. Some values of the hidden curriculum, when unmasked, may become highly contentious. For example, teachers who believe that they run open, thinking classrooms (“I ask lots of questions”) can be shown to be keeping a tight reign on the class through the questions, and promoting a “guess what the teacher thinks” mentality (Hull 1985; Edwards and Mercer 1987; Lemke 1990). The values taught in such classrooms (valuable knowledge is bitty and already known) can run directly contrary to the espoused values of the teacher.

The values of the teacher as expressed in the hidden curriculum need not go unchallenged, even if the students themselves are not explicitly aware of the values. They can be challenged in an equally hidden way, through the misbehaviour of the students. But again, this problem of the breakdown of the educative relationship between teachers and students is not my concern here, important though it is. The point is that much of the value-laden interaction of the classroom is submerged in the lifeworld, and is not made explicit to any of the actors involved.

This places a special responsibility on the teacher. The realization that many of the values being taught in a classroom, even in a community of inquiry, are somewhat hidden from view does not provide a reason to withdraw. We have to act within the horizons of our lifeworld. Nor does it mean that we have to watch every step, to question every action. To try to make every implicit value of the lifeworld explicit at

every moment could only lead to paralysis. Yet teachers must be aware that their own practice is not transparent to them, and that they may not be able to see what others, including their students, can see.

Teachers must take up the responsibility of being in a privileged position in the class both with regard to organizational and epistemic authority. How they wield this power is the question. Is it wielded in the interests of the education of the students, or is it wielded in such a way as to suppress the development of the students both epistemically and morally? Is the teacher also willing to learn and develop? As I argued in §4.141, the teacher must engage in pedagogic action, imposing much on the students through strategic action, without making all the educational goals explicit at all times, but always with the aim of preparing the students for full participation in communicative action and practical discourse. The teacher's actions will be guided by their own virtuous practice, which ought to be well habituated, though this practice must also be the subject of reflection (§5.21). This requires both an inquisitive approach to one's own practice and an openness to critique and questioning from students, but it does not require a withdrawal from the position of classroom authority. The teacher's conduct of the community of inquiry ought also to be regulated by the conditions of the ideal speech situation: not, as we saw in §4.124, by trying to emulate the ideal speech situation in the classroom (which is impossible), but by using the existential and practical aspects to make judgements about the progress of the discussion, and to guard against repression and inequality.

This general survey of the place of the teacher in the community of inquiry has raised a couple of issues with which I shall need to engage. I claimed in the previous subsection that there are conditions under which the teacher ought to challenge the views of students. Given the authority of the teacher identified in this section, it might seem that there are dangers to open inquiry in this recommendation. In §10.23, I will look more closely at this issue. The second issue is related: the teacher's authority (epistemic and organizational) gives the views of the teacher an added weight above those of other students, especially when those views are not explicitly taught, but form part of the hidden curriculum. This opens the teacher up to the charge of indoctrination. I shall now turn my attention to this issue.

9.4 Indoctrination, Ethical Relativism, Moral Indifference and the Path Between Them

Splitter and Sharp (1995, 180) point out the dilemma of ethical education: the twin dangers of indoctrination and ethical relativism. Put simply, they claim that the former is a danger if teachers assert their own moral views strongly, while the latter is a danger if they don't. I think that the situation is considerably more complex in several

ways. Firstly, I shall later develop the position that not all moral views have the same status with regard to indoctrination. In particular, the broad value of respect, as I argued in §6.24, is more basic than many other moral views. Secondly, we have a trilemma rather than a dilemma. If teachers do not assert their own moral views, there can be two outcomes: ethical relativism and moral indifference. By “ethical relativism” here I mean the view that there is no way to differentiate the adequacy of competing moral views, and that hence any view you hold is as good as anybody else’s. Such a moral view can be quite strongly held, well developed and provide moral guidance to a life. Moral indifference, on the other hand, I am using to mark the view that it is not important to hold moral views, and that hence it is a waste of time even to think much about moral matters.

9.41 Indoctrination

On the surface, the danger of indoctrination seems easy to avoid. In the community of inquiry, the questions to be addressed are decided by the students. Once the discussion commences, the teacher remains purely neutral, and thus cannot indoctrinate the students. There are, however, problems with this stance. Firstly, such a stance will be likely to indoctrinate, tacitly and unreflectively, the value of neutrality (I will address this further in the next subsection). But of more immediate interest is the observation that, if the teacher remains completely neutral in all matters of value, then this will include those matters that have to do with the fair and equitable conduct of the discussion itself. Yet, as we have seen, the teacher must take pedagogical action to ensure that the discussion can work, and this requires the enforcement of the values of the community of inquiry itself. These largely spring from that broad value cluster that I have already identified as basic: respect.

9.411 Avoidance of Indoctrination in the Philosophy for Children Literature

Lipman *et al.* (1980, 186) have attempted to address this problem of indoctrination, by drawing attention to

the particular usefulness of the distinction between substantive and procedural considerations with respect to classroom instruction. The teacher... should normally be neutral when moderating discussions among students about specific substantive issues in which value questions predominate. But the teacher in such discussions should definitely be partial to and insistent upon the rules of procedure by which the discussion is carried on. Should these rules happen to become themselves the substance of the discussion, then the teacher should endeavor once again to assume a neutral attitude towards them.

In this analysis, they separate matters of procedure from matters of substance, and advise neutrality towards the latter only. Nevertheless, they recognize two problems here. Firstly, procedural respect cannot be said to be a non-substantive value.

Children might ask why it is that they ought to take turns, or why one with a louder voice should not win out over one with a softer voice, or why the teacher should make the choice as to who will speak next. In this case, the procedure of discussion itself becomes the subject of the discussion. Whenever a procedural rule is questioned, Lipman *et al.* claim, the teacher ought to switch to treating the rule as a substantive matter, become neutral on the matter, and thus avoid indoctrination of the rule.

However, if the distinction between procedural and substantive values cannot hold, then neither can the argument that they can be treated differently. If all procedural matters are in fact at base substantive, and the teacher ought to remain neutral in all substantive matters, then it is contradictory to claim that the teacher has a justification for ignoring this duty of neutrality and pushing a certain value. It is particularly strange to claim this justification holds only on the condition that the value remains unquestioned. I shall return to this problem below.

The second problem is recognized only obliquely by the use of the word “normally” in the second sentence. It hints that there are times when neutrality is not appropriate, but it does not tell us how to identify the circumstances that warrant the dropping of neutrality by the teacher. This question needs more analysis. On page 159, Lipman *et al.* recognise that “the teacher always has the right to intervene and state his or her opinion if the circumstances warrant it.” Only two circumstances are identified for such teacher expressions of opinion: after the other students have already had a chance to respond to the view to which the teacher wishes to respond, or if the community has become strong enough that individual students are able to oppose the teacher’s epistemic authority if they disagree. In my view, these are certainly two circumstances in which the teacher ought to express a view on matters of substance, but there are also others. While teachers must remain aware of their epistemic authority and the potential that their expression of a view might discourage students from exploring the matter further, there are dangers in the other direction, too.

The distinction between procedural and substantive matters calls to mind Habermas’ attempt to ground ethics in a purely procedural moral principle (see §6.24), leaving substantive matters to be decided in practical discourse. In §7.21, I argued the inadequacy of this position, giving reasons why we need to take respect as a basic, presuppositional, substantive moral position. Lipman *et al.*, despite recognizing that there is no sharp cut to be made between procedural and substantive values, advise the teacher to treat discussions of procedural rules as if they were substantive, by taking a neutral stance towards them. If this is merely a tactical neutrality, allowing the other students to mount a defence of the rule, then it is good advice. Certainly, a teacher would be right to allow students to do the work of critique if they are capable of doing

so. However, we may ask whether the teacher ought to maintain this neutrality if the other students prove not to be capable of mounting the critique, or if the class agree to reject the rule, or if the questioner mounts a strong and ingenious case for rejection. The community of inquiry, if it is still in the early stages of construction, may be widely flouting the rules of respect, or incapable of properly defending them. In these and similar cases, I would argue that the teacher must spring to the defence of respect.

The procedural rules of respect are central to the possibility of a community of inquiry. As Reed says, “democratic and egalitarian conditions are non-negotiable prerequisites for philosophy in the classroom. That is, children have the right to have and express their opinions..., we ought to respect other people as persons...” (quoted by Cresswell and Hobson 1994, 29-30). Reed clearly wants and expects these values to be held not only by teachers but also by students. Splitter and Sharp (1995, 36) say “principles of fairness, reciprocity and respect for persons... are fundamental to a community of inquiry” and claim that “the procedures of ethical inquiry are among those which are enacted *and valued* by the community itself” (175, italics added). Philosophy for Children theorists are most certainly *not* neutral towards a certain core set of values that they expect the program to foster in children. It is one of the most ubiquitous features of the Philosophy for Children literature.

This raises the difficult issue of autonomy. The problem becomes acute if children are assumed to be autonomous for, in the light of this assumption, the imposing of values on them is taken to be a violation of that autonomy. However, the account of autonomy that I developed in §3.3 and §4.2 makes it clear that we cannot assume children are autonomous: rather, they are *becoming* autonomous through their interactions with others. They need assistance to achieve a sufficient level of autonomy, and the teacher is charged with bringing this state of affairs about through pedagogical action (§4.141).

If the broad values of respect were to be rejected, Habermas would claim that such a rejection amounts to a performative contradiction. This would be true if the students tried to maintain a discussion while rejecting the values of respect, but they may make the rejection behaviourally, and be content with non-discussion-based educational activities. In situations where the class is incapable of criticizing the rejection, or agrees with it, I maintain that the teacher is well justified in arguing the case for the rule as vigorously as possible and, if this fails and the alternative is to abandon the community, in imposing the rule anyway. As I have argued, teachers are teachers because they have been assigned a role by society - that of education through pedagogical action - and they must use the means, sometimes strategically, that organize the classroom as an educative environment. One of their aims is to induct

students into the capabilities needed to engage in practical discourse, which will lead to the development of their autonomy. The teacher cannot assume that autonomy, reasonableness and other capacities are already in place, and must establish the educational environment in the classroom in which they will be learned.

The standard defence of the community of inquiry from the charge of indoctrination is that it can turn reflexively on any of its assumptions and values (Cam 1994, 19-23; Splitter and Sharp 1995, 180). Certainly, this is its best defence. However, to question reflexively in this way requires a community that is capable of inquiring well into the matter in hand, and it will only be such a community if it already follows certain procedural rules of respect. A young or a dysfunctional community may well not be able to inquire well, and the onus is on the teacher to assist the community (and its members) to grow to the stage where it can inquire well. By the time this stage is reached, the values of respect may be so well in place that students will be much more likely to agree that respect is an essential moral position. Does entrenching respect in this way amount to indoctrination? If it does, then pedagogical action will be indoctrinatory.

If we are to decide whether a teacher's actions in building a community of inquiry through inculcating the values of respect is an instance of indoctrination or not, we clearly require an account of indoctrination. Here is a not untypical example from the Philosophy for Children literature, based on the discussion of indoctrination in Woods and Barrow (1975, Chapter 4).

If students in a community of inquiry were intentionally brought by some non-rational means to the unshakeable belief that freedom of speech, say, is a moral right, then they would have been indoctrinated. (Cam 1994, 36)

In Cam's definition, we have a clear account of the nature of indoctrination. It entails that we can only avoid indoctrination through the rational consideration of values. Rational means, according to Cam, involve the explicit endorsement of a value for well, and freely, considered reasons, while "it would become indoctrination were we to use the practice of free speech to directly inculcate belief in freedom of speech" (Cam 1995, 36). Hence, immersion in practice which does not include the explicit verbal consideration of reasons for and against a value would be an example of non-rational means.

However, it is clear that my account of pedagogic action means that the teacher will quite often take intentional action to bring about a belief in certain procedural values by means that do not involve explicit consideration of the values in question but, rather, by means of immersion in practice. It appears, then, that according to Cam's account

of indoctrination, I am advocating the indoctrination of the basic values of respect. In the next subsection, I will demonstrate that Cam's account of indoctrination is inadequate, being based on a misconstrual of how it is that we are situated in the world, and in §9.413 show that the insistence on respect within the community of inquiry is not indoctrinatory.

9.412 Assumptions Underlying Cam's Account of Indoctrination

Let's acknowledge straight away that many of the procedural values of the community of inquiry do become substantive subjects of the discussion, and that students do become rationally (in Cam's sense) convinced to adopt or retain these values. Clearly, these values have not been indoctrinated. But there are many other values that are never explicitly considered.

I shall take as an example a particular, fairly narrow, procedural value: respecting another's right to finish what they are saying before commencing one's own comment. It may well be that this value has never been explicitly considered by the students, but is nevertheless internalized so that they come to think, unshakeably, of "butting in" as rude; something that ought not to be done (even though they might sometimes slip). For the sake of argument, I will assume that, for the students in question, a number of the crucial ways in which this value has been inculcated have been due to the intentional actions of the teacher, who has shushed people who tried to butt in, or waved them down, or sometimes even said "Wait your turn" and so on. Through a Vygotskian process, the students have internalized an aspect of communicative action that they have engaged in within the community. Explicit verbal reasons for this rule have neither been advanced nor considered. Rather, this (non-explicit) value that the students have internalized (through immersion in practice) forms part of their lifeworld, taken-for-granted, held unshakeably, and not present to them consciously. Is this a case of indoctrination?

Now, we must concede that, since the process by which this value was inculcated is non-rational in Cam's sense, it is, by his definition, a case of indoctrination. However, underlying Cam's definition are two assumptions, which I have shown in previous chapters to be untenable. The first assumption concerns the way in which persons are situated in the world (§3.124), and the second, which arises from the first, concerns the nature of belief acquisition processes, and how they justify the beliefs acquired (§3.23). Under Cam's assumptions, it is impossible to explain how it is that we become able to operate within the world, and most of our beliefs and values would be unjustified.

In Cam's account, persons are already rational, autonomous beings who are presented with a smorgasbord of values. Ideally, from their value-neutral position, they rationally assess the values that surround them and choose which to adopt. Sometimes, however, manipulative others trick or coerce them through non-rational means to adopt other values. This picture is derived from the philosophy of the subject, as discussed in §4.141.

The contrary view, which I discussed in §3.124, is that we are always already immersed in the context of the lifeworld, and that coming to be able to operate within the lifeworld cannot be a matter of explicitly and rationally (in Cam's sense) evaluating the context. In that section, I drew on the arguments of Heidegger (1962), Polanyi (1958), Haugeland (1985), Dreyfus (1992), Cosmides and Tooby (1994), Clark (1998), and Winograd and Flores (1986). Many of our values reside, in Heidegger's phrase, in our pre-reflective background and form, in Gadamer's (1975) phrase, part of our prejudices. As he says, "All understanding inevitably involves some prejudice... 'Prejudice' means a judgement that is given before all the elements that determine a situation have been finally examined" (239-240). Much of our thinking is not reflective and we cannot be conscious of all the thinking that we do. When we are immersed in a practice, we do not explicitly consider even a good deal of the potentially conscious thinking that we are doing, especially when some other topic is engaging our attention. One consequence is that some of the values we exhibit in our actions are inconsistent with our espoused values, while others are quite consistent. But without the presupposition of this horizon, including the presupposed values, it would be impossible to make any sense of the world or to operate within it.

Within the lifeworld, our learning is Vygotskian (§3.23 and §8.23), so that we come to think and value through engagement in socially constructed thinking and valuing. Some of this is explicitly evaluated, some is not, but the ability to evaluate both the thinking and the valuing is gradually constructed. Since our ability to evaluate explicitly is weakest when we are youngest, and only gradually strengthens, many of our values are acquired well before they could possibly be fully evaluated. However, as I argued in §4.22, in order to become an ethically sensitive person, we need to develop not just values, but also reasonableness about values. In Aristotelian terms, we can say that reasonableness is needed not only so that we can apply the values we have in the right way, but it is also needed to be able to exercise care about the sort of person we are. Consequently, to be able to become a virtuous person (both morally and intellectually) requires that certain values, and capacities of reasonableness, are inculcated within us through non-reflective means. These may, or may not, be subsequently endorsed (as we shall see), but they cannot be initially acquired through

explicit choice. This fact about how these values and capacities have been acquired carries, in itself, *no* implication as to whether they have, or have not, been indoctrinated.

This leads to Cam's second assumption concerning the nature of the fixation of belief. Cam's idea of rationality is heavily dependent on what I have labelled the critical aspect of reasonableness in §3.121. He says

we need to be not only aware of the values that we promote in the classroom... but also careful to see that these values are subjected to scrutiny by the community. This is not just a matter of coming up with a list of reasons to support these values, but of subjecting them to genuine assessment. (ibid, 35)

Such rationality requires recourse to argumentation backed by reasons and conducted in language (see Habermas' account of argumentation outlined in §4.123). This, in Cam's view, is the only legitimate way to acquire values. Values can be acquired in other ways, but until they have been rationally endorsed, they have merely been indoctrinated into us. Cam admits that "the stand against indoctrination is... difficult to achieve in practice," partly because of this requirement to subject all beliefs to scrutiny. Further, he claims that we are "all of us working against a background of indoctrination, including the influence of the school, the home, peer group, televisions and so on" (ibid., 35). The problem raised by this account of indoctrination is that it labels much of the ordinary activity of the lifeworld indoctrination. If this is so, then indoctrination will be difficult to avoid indeed, even in the community of inquiry.

I will show that, while the process of gaining a value through immersion in practice is (in Cam's sense) non-rational, it is nevertheless not, solely for that reason, a case of indoctrination. I will outline the conditions under which the charge of indoctrination is not justified.

9.413 An Alternative Account of Indoctrination

I have argued that to say that a value has been acquired in the lifeworld (non-rationally, in Cam's sense) is not to say that it has been indoctrinated. But, of course, it is not to say that it was not indoctrinated either. Suppose I acquire in this manner a belief in the supremacy of my race over others. Subsequently, I recognize this to be an immoral value, inconsistent with other values I have acquired reflectively and hold to be more basic, so I change it. Then I would say that my old belief in racial superiority had been indoctrinated into me. This example points to a counterfactual account of the nature of indoctrination. Whether a belief has or has not been indoctrinated depends on what my explicit judgement about my acquisition of it *would* be, *were* I to make such a judgement. I shall expand upon this view.

My claim is that we are always already situated in the lifeworld, before we can reflectively evaluate any of the beliefs and values that we have. Many of the beliefs and values we hold are inculcated through this immersion. Nevertheless, through our immersion, we learn in Vygotskian fashion not just beliefs and values, but also reasonableness. If we were not immersed in the lifeworld, then we would not be able to learn, and teaching would itself become impossible. Through our engagement in social reflection on matters of belief and value, we become gradually more able to reflect ourselves. As we become more able to reflect, we are then (in Aristotelian terms) gradually able to take on the responsibility for our own characters, and begin to look reflectively at the values we always already have. We are becoming not only habituated to the virtues, but also to reasonableness, without which we cannot be virtuous.

In this account, indoctrination cannot be identified with this lifeworld process by which we come to hold a value, since we cannot escape acquiring values through non-rational (Cam's sense) means. Rather, a value is indoctrinated if it meets one or both of the following counterfactual conditions.

- A value has been indoctrinated if, were the value to be subjected to reflective evaluation by the person in whom it has been inculcated, it would not be adopted;
- A value has been indoctrinated if, were the means by which the value was acquired subjected to reflective evaluation by the person in whom it has been inculcated, those means would not be endorsed.

The corollary of these conditions is that a value may well have been acquired by non-rational means (in Cam's sense) but not have been indoctrinated provided that, were the holder of that value to reflect on both the value itself and the means of acquisition, then both the value and the means would be approved.

I shall expand and illustrate my account by reference to the value of respect, in terms based on the Aristotelian/Habermasian account of moral development (see §7.2). Recall that both Aristotle and Habermas assert that children come to moral discussions already holding some values. This I have described above using Habermasian terminology. The child is always already situated in the lifeworld: a world which is intersubjectively created within the physical environment by embodied and already socialized creatures. The lifeworld provides facts and values that are taken-for-granted and acted upon quite unreflectively. As the child becomes more able to reason and to question (and we have to bear in mind that these capacities are built dialogically), some

of these facts or values emerge from the lifeworld and become thematized. They can then become the substance of discourse.

Communication is seen as a comprehensive goal or value: by attempting to communicate, actors in the social world demonstrate that they value the ability to communicate with others. The community of inquiry arises out of the lifeworld in which such communication can take place, and continually draws upon it. Without this overlapping and shared background, it could not get started. Habermas, as we have seen, asserts that amongst the presuppositions that must be in place in the lifeworld are those that underlie the possibility of communicative action at all: the presuppositions of respect. This is not to say that the rules of respect (such as Alexy's list presented in §6.1) are known and agreed to explicitly by all participants: far from it. Yet when communicative action takes place, they are acted upon locally, and any local violation is seen as disruptive to communication: it is not valued. This inculcation of the local variants of the rules takes place with all speakers as they develop the ability to enter, however haltingly, into communicative interaction with others.

In saying this, I recognize that the presuppositional rules of communication (closely related to what I have also called the broad value-cluster of respect) are not purely procedural values. They are substantive values that are very often treated merely as procedural values. Only when they are problematized does their substantive nature become apparent: this is probably true of all the values of the lifeworld. But some inkling of them is present in every communicative exchange, and the infant that is learning to communicate is learning these values, unconsciously for the most part. This is not to say that they cannot be made explicit or questioned, for they often are. Admonitions for the young child to listen, or not to speak over the top of someone else, are common, and children are not backward in asking why. Thus lifeworld values can be part-thematized and part-redeemed in discourse at a level accessible to the developing communicative subject.

The community of inquiry self-consciously places a good deal of emphasis on explicitness: much more than the lifeworld does. It is a structured situation, in which one person - the teacher - has explicit organizational responsibility, and part of this responsibility is precisely to build the possibility of, and to encourage, explicitness. The lifeworld does not have such formative guidance, but is jointly constructed by all, some of whom have greater or lesser organizational authority in various fragments of the lifeworld.

When students become part of the community of inquiry, they bring all their lifeworld values with them. Some have never been brought to conscious attention, but many will have been, at least to some extent. Some of those to do with respect are very likely to have been made explicit on many occasions, both before entering school and certainly within school. Just how much conscious attention has been paid, how much these core values influence behaviour and just what detailed content they have will vary from student to student. We might capture this by saying that each of them will have a different respect-horizon. To the extent that these overlap, and overlap with the respect-horizon of the teacher, the students will unproblematically practice respect in the community of inquiry. When they do not overlap, especially when one student's actions jar with the teacher's respect-horizon, the teacher can, verbally or by non-verbal means, require that the student comply with the teacher's value. One of two things can happen at this juncture. Either the student will acquiesce, possibly modifying somewhat their corresponding value, or the student can question (verbally or behaviourally) the value that jars.

The first of these possibilities is one of the major processes by which the lifeworld values are built in the first place (possibly, as I intimated in Chapter Two, on an empathetic genetic base). We must bear in mind that the students are becoming-persons, not finished persons (as, to a lesser extent, are adults). They are continually learning, Vygotsky-style, the facts, values, concepts and capacities that are collectively constructed in the human interactions in which they are embroiled. Yet, insofar as they have already internalized the capacity to reason and to question, they can also follow the second possibility. They can problematize the value. Then, insofar as the community of the students can jointly explore the value, the teacher ought to confine input to procedural matters. But equally, insofar as the students are not yet capable of exploring it well, the teacher must intervene, both procedurally and substantively, so as to bring about the conditions under which the community of students will be able to enter as full a practical discourse as possible. Since this practical discourse will be in the joint zone of proximal development of the students, they will benefit by being able to internalize the capacities, concepts and (most importantly in this context) values that they jointly construct with the teacher's scaffolded assistance. Here is another of the processes for lifeworld learning, only it has been sharpened and deliberately wielded by the teacher.

In this account, I have hinted at, but not yet made explicit, an account in which some values, or value-clusters, stand in a higher or more overarching position with respect to other values. To expand on this, I shall draw on Raz's (1986, 289 ff.) discussion of the nature of our valuable goals, and their relation to our well-being and our

morality. He claims that our well-being requires that we have goals or values, some of which (like the value attributed to eating) are mainly biological, others of which we have adopted deliberately, and still others of which we have drifted into. This last category covers those that we acquire through immersion in a practice, for he says that our values are not always come to by explicit reasoning. Importantly, he points out that our values and goals are nested hierarchically. The highest level values - those he calls comprehensive goals - pervade our life, while the lower level values have little weight, and are subsumed under our comprehensive goals. These, he says, are often not means to the end of the larger goal, but constitutive elements of it.

Our highest values and comprehensive goals, Raz claims,

are too dense to allow explicit description or learning, they can be learnt only by experience, direct or derived (e.g. from fiction). It is if course not only the learning which is not explicit. Even once the patterns of behaviour have been learnt much, indeed most, of our behaviour remains based on learnt semi-automatic responses (i.e. ones which we can, usually with some effort, suppress, but which we normally do not deliberate on and which we are not explicitly aware of). (312)

Raz is claiming that individuals can only acquire and maintain their goals through familiarity with social forms; that is, through immersion in practice in the lifeworld, or a surrogate for it like literature. The high-level values that are learned in such a way, he says, are expressed through a dense series of behavioural responses, each representing a lower level value that is constitutive of an overarching comprehensive value. These lower level values are enacted and implicit rather than expressed and explicit, but they are subject to a degree of control and can be made the subject of consideration if necessary. In this account, Raz agrees with the one I have developed above, but he adds to it a distinction between levels of values.

The pervasive, comprehensive values to which Raz refers are more likely to come to awareness simply because of their pervasiveness. We are more likely to be conscious of the fact that we value communication, and the broad value of respect for our partners in communication, than we are of, for example, the “thousand tiny clues of what is known as body language [which] contribute... to the success of the developing relationship” (ibid, 312). Yet each pause in conversation, or encouraging eye movement, or whatever, instantiates a value of respect for the other.

To say that we ought to bring to awareness and evaluate all the lower level values that we hold, and act out, would be to call for paralysis. It would make the world we are immersed in strange to us and alienate us from it. This is not to say that they cannot be considered, because circumstances (such as unexpected consequences) may make one of them strange to us, and cause us to deliberate.

Yet, as I have noted, the more pervasive values are more likely to be opened up for deliberation, as they are involved in our most important choices and actions. In many situations, the choices we make and actions we take are unreflective, even when they involve comprehensive values, but there will be situations where several values clash, or where the situation is sufficiently novel to cause us to wonder how best to instantiate our values. On such occasions, the comprehensive values emerge from Polanyi's subsidiary awareness and into focal awareness (Polanyi 1958, 57, see §3.1242). Since it is these comprehensive values that guide our lives, then these are the values that most need to be open to reflective evaluation.

However, to endorse the engagement in reflective evaluation when appropriate is in itself a comprehensive value, instantiating a large number of lower level values. Yet it stands in a peculiar relation to the other comprehensive values that a person may hold, because it is a necessary value for the very possibility of testing for indoctrination. It is thus a value that underwrites any anti-indoctrination value.

This discussion of the hierarchical structure of value systems points to a modification of the account I gave earlier of indoctrination. In that account, I claimed that the decision on a charge of indoctrination was a counterfactual one. An actual decision as to whether such and such a method of inculcating such and such a value is, in fact, indoctrination, turns on actually subjecting the method and the value to scrutiny. This in turn depends on the allegedly indoctrinated person being capable of carrying out the scrutiny. It would thus seem that a society could avoid the charge of being indoctrinatory by inculcating comprehensive values that discourage puzzlement and questioning, and by not inculcating the capabilities of inquiry, thus avoiding the test of indoctrination as much as possible.

So, we need to add another condition: not at the value level this time, but at the social.

- A society (or subsection of society) is indoctrinatory if it does not provide its members with the wherewithal and the opportunity to reflectively evaluate values that become problematized.

Here, then is an account of indoctrination and its avoidance. Societies, through the social actors that make up those societies, will inevitably create the conditions under which children will absorb the many values of the lifeworld - values which are not necessarily coherent. The processes by which this happens can be either explicit or implicit. This is not just inevitable but also necessary, for action must arise in and from a rich lifeworld. Teachers are not exempt from this, and the community of

inquiry, like any social action, also inculcates values. To apply the word “indoctrination” to this makes no sense at all.

However, if the society (or subsection of society such as a family, school, classroom etc.) does not provide both the wherewithal and the opportunities to problematize the values that form part of the lifeworld(s) of its members, then it can be accused of being indoctrinatory. It is not so much the inculcation of values or the insistence upon their observance which makes for indoctrination, but the denial of the chance to make those values the subject of a practical discourse. The community of ethical inquiry is not indoctrinatory because it does provide both the wherewithal (in that it promotes the development of reasonableness) and the opportunity (in that it allows and encourages the problematization of values when those values seem to need to be problematized). Thus, we cannot say that a child has been indoctrinated with a particular value solely on the grounds that the child has never explicitly rationally endorsed it. Indeed such a total stocktake of all our values may be extraordinarily difficult or even impossible to achieve. Nor can we say a value is indoctrinated if, on one occasion, it is inculcated without explicit rational endorsement. Indoctrination is a charge which must be brought on the grounds of a total approach to a group of people over time, where these people are denied the wherewithal and/or the opportunity to submit their values, when questioned, to practical discourse.

On this account, then, we do not need to call for teachers to remain neutral about matters of value so as to avoid the charge of indoctrination. Nevertheless, we can say that there are good pragmatic reasons why teachers ought to be wary of asserting their values, and that there are important distinctions between types of values which ought not to be asserted strongly, and those for which (with appropriate caveats) strong assertion is acceptable or even desirable. I mentioned these at the start of this subsection. There is a distinction, rough though it may be, between those procedural values necessary for the enterprise of communication, and those substantive values about which communication can take place. The former may be converted into the latter, when it seems appropriate, and their exact nature can be clarified, but they are still needed in some form to create the conditions under which discourse about values is even possible. This possibility is the insurance against indoctrination. The epistemic authority of the teacher is strong, and hence the teacher ought to be wary, at least until the students in the class have attained the ability and inclination to test the teacher’s views in the same way they can each other’s, of asserting substantive values which are not connected to the development of discursive capabilities.

In summary, the charge of indoctrination has often been understood in terms that rest on an (often implicit) mistaken assumption that children already have reasonableness

and autonomy. If this were so, then the demand that all values must be rationally chosen and endorsed might make some sense. But when we consider that children come into the world with certain predispositions (§2), which are then supplemented by much knowledge and values that are implicitly acquired as background through immersion in practice (§3.124 and §3.125), we begin to recognize that rational endorsement can only be expected once many other capacities have been developed to a sufficient degree. Becoming virtuous requires an Aristotelian habituation to the virtues (§5.2), or (equivalently) a Habermasian immersion in the lifeworld (§6.1). Yet this is not a claim that reasonableness (or rationality) has no part to play in moral development. Aristotelian practical wisdom (Habermasian practical discourse) must be exercised in the process of habituation, and reasonableness itself must be habituated (§5.22). Only through this cycle of recursive habituation and reflection can the capabilities to make reasonable, autonomous decisions about values be built (§4.22). Building them is (to an important degree) the role of the teacher, who must exercise pedagogical action to bring their development about (§4.141). The consequent inculcation of values in the classroom, like that in the lifeworld, is not indoctrinatory provided the three conditions I have advanced in this section are met: students would endorse the values once they are reasonable and autonomous to do so; students would endorse the means of their inculcation; and the classroom provides the opportunity and wherewithal to consider such endorsement, if the students so choose.

9.42 Ethical Relativism and Moral Indifference

It is the fact that we live in a pluralist liberal democracy, rather than in a closed, monocultural one, that has made the dangers of indoctrination loom large. In a diverse society, to try to force one set of views on another, particularly when those views run counter to the beliefs of a minority or disadvantaged group to which the other belongs, has been seen as violating important rights of the other. In response to such worries, great efforts have been made in schools to avoid indoctrination. It is my belief that too great an effort in this direction has dangers of its own. Above, I identified two: ethical relativism and moral indifference.

Since there are many varieties of relativism, I need to be clear about the sense in which I am using “ethical relativism” here. If I come to believe that moral values are purely personal, not the sort of things that are subject to outside criticism, and that hence any view I hold must be as good as any view anyone else holds, then I am, in this sense, an ethical relativist. This is a deeply individualistic view about the criticizability of moral views. I explicitly distinguish this type of ethical relativism from the suite of philosophical theories that go under that name. A relativist meta-ethical theory may be tenable, but only after philosophical consideration which must take place in a

community of philosophers. Most well-thought-out relativist meta-ethics do allow for criticizability. And any theory that relativizes morality to a society is incompatible with the sort of ethical relativism I am picking out here.

Moral indifference is the other danger. I am morally indifferent if I believe that it is not important whether I am for or against abortion, war, violence in the media, sex before marriage, or whether I am personally honest, or take account of the effects of my actions on others and so on. Moral indifference means giving such commitments a very low weighting in the scheme of things. If moral views are thought to be unimportant, then the disposition to take morally good action when in a position of temptation will also be low. Nor will consistency of action matter much. Such persons may be good calculators of their own personal advantage, like Hume's (1955, 236) sensible knave who "observes the general rule, and takes advantage of all the exceptions." Or they may, lacking strong values, be prey to the urgings of those that surround them.

If it is recognized that values are important, and that students ought to have values, but there is a strong commitment that these values ought not to be indoctrinated, then we might conclude that any offering of views, or evaluation of students' views (especially by teachers) is unacceptable. There are two well known values education programs that take this approach: Values Clarification from the US (Read and Simon 1975; Rath, Harmin et al. 1978); and the Schools Council Nuffield Humanities Project (1970) in the UK. Both programs are based around the consideration of values under the neutral eye of the teacher, though there are important differences between them.

In the Humanities Project, the teacher is expected to act as a "neutral chairman" who "regards it as part of his responsibility not to promote his own view" in class discussions concerning controversial moral social issues, but "the teacher as chairman of the discussion should have responsibility for quality and standards in learning" (1970, 1). This description of the teacher's role relies on a strong content/process distinction, which we have seen is not tenable. However, we have also seen that some sort of distinction can be made between promoting values that have to do with the conduct of the discussion, and promoting values that have to do with clearly controversial content.

In discussing this Project, Straughan (1988) quotes Warnock's criticism that "children must not be deprived of the spectacle of a teacher who holds, and clearly expresses, moral views." I agree, for the danger is that to model neutrality is to convey the message that neutrality on moral issues is a good thing - the danger of moral indifference. But we need to find a balance that also recognizes the epistemic authority

of the teacher and the coercive weight this can carry. Teachers can and ought to hold and clearly express views, provided that they both make it clear that they have a responsibility to open these views to critique, and that they provide the students with the wherewithal and opportunity to engage in that critique. Despite these comments, the approach suggested by the Project, in which the children are encouraged to criticize and defend moral positions in class discussion, does resemble the community of inquiry approach much more closely than Values Clarification, in that it recognizes the need for public assertion and defence of moral views.

The attitude of Values Clarification to values education is much more individualistic. Thus, Rathes *et al.* (1978) say that “values are personal things” (34), and that “something will not qualify as a value if... it has not been *freely* chosen (there is no room in this theory for values that are imposed by outside pressures)” (47). It is easy to see the Kantian roots of this view. Children are autonomous, rational individual agents who are capable of making a free choice. Choices are personal, and ought not to be subjected to outside pressures. There are, however, non-Kantian elements as well: our inclinations and desires are allowed to influence our choice.

Unlike either Philosophy for Children or the Humanities Project, little role is seen for the social negotiation of values. Rathes, Harmin and Simon (in Read and Simon 1975) say:

Choices cannot be considered sufficiently free if each one is to be weighed, approved, disapproved or graded by someone in charge. (78)

This means that we are dealing with an area that isn't a matter of proof or consensus. It is a matter of experience.... It is important to note that our definition of values and valuing leads to a conception of these words that is highly personal. It follows that if we are to respect a person's life, we must respect his experience and his right to help in examining it for values. (80)

We are interested in the processes that are going on. We are not much interested in identifying the values which children hold. (81)

While the first sentence precludes the teacher from making any evaluation of the student's choice, it is clear from the disapproval of proof and consensus in the second that it is not just authority figures, but everyone, who are excluded from evaluation. The source of values is extremely unclear. On the one hand, they are freely chosen from alternatives, whereas on the other, they are the result of experiences we have. There is certainly no room for the development of the capacity to choose in this picture: it is present all along. Finally, there is again a very sharp process/content distinction. The actual values held are not seen to be of any great importance. Ethical relativism, indexed to the individual, is seen as a positive outcome.

One of the major problems with the Values Clarification approach is that it makes no effective distinction between moral and other values. Whether one values blue clothes over red clothes is seen as being much the same as whether one values cruelty to animals over kindness. To aim simply for clarification in matters of taste is unobjectionable, though it must be pointed out that this still avoids the question as to whether taste can be educated. However, when it comes to matters of morality, a non-judgemental acceptance of certain extreme positions is in itself immoral.

This is not to say that clarifying the values that students hold is not an important part of moral education. It is difficult to engage in moral discourse if one's values are not clear to one. However, to just ask students to clarify their values and not to engage with others in evaluating and defending them, possibly even changing them for good reasons, is to risk ethical relativism of the most thoughtless kind. As Fraenkel (1977) comments:

Values clarification does not help students to appraise critically their own or anyone else's views. Rather, it encourages them to accept uncritically the values of their society. In fact, it teaches that one value is as good as any other.... Values clarification does not provide student with any way to deal with the internal (and often external) conflict and uneasiness... opposing values produce. (47)

Harmin and Simon (in Read and Simon 1975, 407) recommend small group discussions, but the only rationale for this they offer is to "allow everyone to get his two cents in." This is not the way to build the capabilities for critical appraisal of values. Nor, as Fraenkel points out, does it help students to weigh up the imperatives of clashing values, in situations where, say, the value of honesty demands a different reaction from the value of kindness. Clarifying a value gives little help in using it in specific situations.

Now I shall turn to the position that moral education is no part of the school's brief. As I discussed in Chapter One, this view holds that moral education, or training, comes under the purview of the family and/or the church and/or some other institution. Schools (this view is particularly expressed in relation to government schools) are meant to be value neutral and not to promote any one moral code over another. We have seen that this is an unachievable aim, for schools promote, and must promote, some moral values if they are to work at all.

Underlying this position is the view that schools ought not to promote partisan moral views on matters under current contention in society, and that therefore they ought not to raise such issues at all. But it is certainly possible to agree with the premise and dismiss the conclusion. If schools studiously avoid any engagement in any contentious social issue at all, especially ones in which students are likely to be

involved, then the implicit message is that such disagreements are not of any great value in the schools', or the teachers', eyes. While this may not be true at all - the teachers may individually and severally hold strong views on the matters - without any evidence to the contrary, this is the conclusion students are likely to draw, and may even internalize. This stance thus risks promoting moral indifference. This danger is the more acute if no other important element in a student's life is itself addressing such moral views.

9.43 The Path Between

Many moral matters are contentious: some, within our society and some between our and other societies. Many other moral matters are not contentious, but this is not necessarily a guarantee that we have come to a correct moral evaluation of them, as consideration of the widespread acceptance of slavery at one time shows. I have identified three dangers that can arise from attempts to morally educate students in the face of this diversity and uncertainty.

The fear of illegitimately imposing views on students lies behind modern educational attempts to avoid indoctrination. Such a concern was not particularly evident when society was more uniform with regard to moral views. However, I have argued that these attempts to avoid indoctrination, important as they are, face the opposing dangers of ethical relativism and moral indifference. To avoid these, teachers need to help students to see that moral values are important, that they are held for reasons and that they can be defended against attack, or abandoned if the opposing reasons are good enough. Teachers need to demonstrate that they are committed to their own moral beliefs and are prepared to defend them thoughtfully, if they want students to turn out likewise. Nevertheless, teachers also need to be aware of the dangers that their epistemic authority will lead to students adopting their beliefs thoughtlessly.

To walk the line between indoctrination and the encouragement of ethical relativism or moral indifference requires nuanced and situated judgement from the teacher. It is the open invitation to question, implicit in the whole community, that avoids the charge of indoctrination. Equally, it is the assumption that no view is immune from criticism, nor without reasonable defence against the criticism, that avoids ethical relativism. And it is the commitment to questioning and reasoned critique that avoids moral indifference.

9.5 Towards Detailed Practice

In Chapter Nine, I have taken a broad brush look at the place of the community of inquiry in the classroom. After initially making clear just exactly what I mean by the

phrase “community of inquiry,” I addressed the reasons why such an approach might be expected to be a highly effective way of developing reasonableness in §9.1. In §9.2, I turned to the development of ethical persons, and showed how the community of inquiry meshes with the virtue/discourse meta-ethical theory developed in Part II. I showed that the community of ethical inquiry has the resources to avoid two weaknesses sometimes identified in overly rationalistic approaches to moral education: the lack of linkage of judgement to action; and the use of moral reasoning to engage in moral rationalization. My focus changed somewhat to the role of the teacher in §9.3, where I emphasized the delicate balance between the teacher’s authority and the development of children’s autonomy and reasonableness: a balance that requires nuanced pedagogical judgement and action. Finally, in §9.4, I looked more carefully at one of the major implications of the considerations of §9.3: the possibilities of indoctrination, ethical relativism and moral indifference that spring from various misunderstandings of the teacher’s role.

These broad brush accounts will, I hope, be of some assistance to classroom teachers who wish to run communities of ethical inquiry in their classrooms. However, they do lack some of the detail that teachers would be interested in. Chapter Ten will address the linking of my philosophical account from Parts I and II to the day-to-day organization of a community of ethical inquiry.

Chapter Ten: Practical Implications

What are the practical implications of this discussion for teachers trying to assist their students to develop morally? This is possibly the most important question to be answered, for all the philosophical investigations I have carried out will be fruitless unless they offer guidance to the teacher in the classroom. In a phrase, the answer is simple: teachers ought to engage their students in a community of ethical inquiry. But such an answer is uninformative until I link the philosophical conclusions I have made to the specific features of the community of ethical inquiry.

The community of ethical inquiry is a specific instance of the community of inquiry (see §9), in which the core content to be discussed concerns ethical and moral matters. As we saw there, such a community has five stages: the offering of the text; the construction of the agenda; solidifying of the community; using exercises and discussion plans; encouraging further responses. In this section, I will briefly reflect on the formation of the community of inquiry, then consider the implications of the thesis so far for each of these stages in turn, before finally stepping back to look at the wider picture. The focus of this discussion will be the teacher, and guiding it will be the notion developed in §4.141: pedagogic action.

Pedagogic action, it will be recalled, is an indissoluble melding of open strategic action with communicative action and discourse, used by a more experienced person interacting with a less experienced person with the specific aim of developing the latter's competencies, particularly those needed to engage in communicative action and discourse. These latter competencies, it will be remembered, include reasonableness and communicative autonomy. What justifies this action is precisely the facts that the less experienced person or persons do not yet have the fully developed capacities, and that the more experienced person has a responsibility (in a role such as parent or teacher) to develop them.

For teachers, the open strategic action includes such activities as planning for lessons, requiring certain behaviours from students, intervening into student activities in ways which may appear to the students to have one aim, but also have another (pedagogic) aim and so on. Choosing just which strategic action, at what particular time and in which circumstances, will assist in achieving pedagogic aims is an important part of the teacher's professional competence.

The first pedagogical act that teachers must engage in, once they decide that the community of inquiry is a suitable educational methodology for their class, is to constitute the community. As we saw in §8.21, a community can be tight or loose, but it needs to have some commonality. Communities can be constituted in a number of ways: for example, some communities form because their members jointly decide to form them, while people are born into other types of communities. In the case of the community of inquiry in the classroom, the community is formed solely at the behest of the teacher. In similar fashion to the community of the class itself, the students get little or no say in its formation. This is not to say that the community of inquiry is entered unwillingly by all students, although this will most likely be the case with some. Most students do have some commitment to education, and place their trust in the teacher to provide it (§4.14).

However this may be, students do not have a say in the fact that they are being made members of a community of inquiry initially. Yet, if the community is to work, they have to decide to be an active member of that community, and this requires, as we saw in §8.21, some commonalities. An interest in doing as the teacher asks, and quite likely also an interest in becoming educated, are obvious first starting points, but as the community of inquiry develops, other interests ought to come to the fore. Chief amongst these, in a successful community, will be the interest in inquiry itself - inquiry into the concepts and puzzles of the community and inquiry into each other. A growing community of inquiry will become better able to inquire, as the members' reasonableness and autonomy develop. Allied to this development will be the growth of a trust in each other as fellow inquirers, and an awareness that inquiring well for oneself is dependent on being able to inquire well with others (§3.241). Part of that reliance on others will arise from the recognition and valuing of diversity within the community: diversity of views, of learning approaches, of styles of reasonableness. And finally, there may well come a recognition that the boundaries of the classroom community of inquiry are porous. Membership of this community (unlike some) does not exclude outsiders, or interaction with others.

There is much more that could be said on the matter of the constitution and development of the community of inquiry, and even the possibility that a community might, under some conditions, start to break down and decay. Certainly, the community will evolve over time, and we cannot expect that it will exhibit its full power right from the start. As it evolves, the teacher will need to keep re-evaluating the appropriate ways in which to wield pedagogic power, for greater input, of a different character, will be needed for a newer community than for a well-established one. This is a matter for the professional judgement of the teacher. However, I will

not go any deeper into the formation and evolution of communities of inquiry here. I will turn instead to the already constituted community of inquiry.

Under the general guidance of the idea of pedagogical action, I will consider the three main concepts that I have developed in this thesis: reasonableness (outlined in §3.1); communicative autonomy (summarized in §4.22); and the virtue ethics/discourse ethics account of moral development (outlined in Chapter Seven). To structure this discussion, I will consider, in §10.1, the preliminary phase of the community of inquiry - the offering of the text and the construction of the agenda (Lipman's stages I and II) - and then in §10.2 the discussion phase (Lipman's stages III, IV and V).

10.1 The Preliminary Phase of the Community of Inquiry

Let's turn then to the first two of the five stages of the community of inquiry: the offering of the text and the construction of the agenda (see §9). These form the preliminary phase in the sense that they provide the platform from which the dialogical inquiry can be launched. As can be seen from Lipman's phrase "setting the agenda," one of the major premises of the Philosophy for Children program is that the students drive the agenda. This is the advice offered in one of the teacher's manuals: "*Your aim is to work with what the students themselves find interesting*, rather than to set the agenda yourself" (Lipman, Sharp et al. 1984, ii, italics in original).

Yet the agenda is constrained by the text that is offered. To talk of an offering of the text is perhaps to elide two important questions: who chooses to offer the text, and on what grounds do they choose (or write) just *this* text? The answer to the first is fairly obviously the teacher, but this fairly obvious answer is somewhat misleading. Teachers can choose the text in two senses. Firstly, they can choose an already collected set of Philosophy for Children texts, such as one of the Lipman novels (e.g., the ethical inquiry oriented Lisa, Lipman 1983), or another purpose written set of stories (e.g., Cam 1997), or a commercial picture book chosen from a list in a Philosophy for Children oriented program (e.g., Sprod 1993). Alternatively, they can identify or write a text themselves. Clearly, the possibilities available to the students in setting the agenda will be somewhat curtailed by this choice: I shall return to this later.

While, in the first case, teachers in a sense choose the whole collection, they are taking on trust to some extent the choices made by the writer or editor of the collection. Teachers judge, or are assured, that the texts are suitable for the purpose of initiating the community of inquiry, but they are unlikely to look closely at the specific features of the texts which make them so. This job has been done by the writer or editor. In the second case, teachers take on (or ideally ought to take on) the responsibility borne

by the writer or editor in the first case. They ought to be aware of the grounds that underlie the choice of a particular text, or way to construct a suitable text. There are a number of grounds that arise from my discussion above.

My account of reasonableness suggests that the stories chosen to be the triggers for a community of inquiry need to have a very rich narrative structure. Kohlberg's dilemmas are sketchy, lacking much in the way of contextualization and human character development, because they rely on a picture of moral reasoning which, being confined to rational moral judgement based on principles, is quite narrow. An appropriate text to offer to a community of ethical inquiry needs to contain the resources to allow engagement in not just thinking that is heavily biased to the critical aspect (important though this is), but also to the other aspects.

The complex situations depicted in literary narratives provide contextualization for the events and reactions of the characters, and also provide numerous connections to the contexts of the lives of the readers, through analogy and anecdotal similarities, all of which will exercise the contextual aspect of thinking. Indeed, as the specific contexts of the texts change throughout the narrative, or from one narrative to another, yet the same or similar cognitive tools and concepts reappear in these different contexts, the conditions for transfer across contexts are created, especially when these tools and concepts are addressed meta-cognitively.

A rich literary text contains invitations to imaginatively enter the text, extending through the creative aspect to thinking the background lifeworld of the characters, which can be explored collaboratively in discussion. Furthermore, the imaginative structures of moral thinking identified by Johnson (1993), such as the metaphorical extension of prototypical moral concepts to neighbouring ethical situations, can be practiced as this creative filling in of the context by different participants in the discussion produces slightly different scenarios. As we try to decide whether to call what a character said a lie, for example, different students might ask us to consider what our reaction would be if the character had himself been told the falsehood and (perhaps foolishly) believed it, or had been intending to do good by telling it, or had meant to be funny in saying it.

Fleshed out characters have emotional reactions to the events in which they are involved. They act within their social community according to their commitments. In the Kohlbergian dilemmas, respondents must factor out these emotions in order to be assessed as being at the higher stages. Gilligan (1982, 100-103) discusses the tendency for female respondents to draw on what I have called the committed aspect of thinking, resulting in their being assessed as less morally developed. Stories in which

the emotional aspects of the characters' lives are portrayed enable participants in the community of ethical inquiry to engage with the committed aspect of thinking.

Well drawn characters in a rich narrative are not disembodied presences; rather, they think and act as a whole person, mind and body. Their actions can often be said to be "thinking with their body" and they have visceral as well as intellectual reactions. Their embodiment influences their actions, thoughts and situation in the world. These are features which cannot be captured in sketchy scenarios, and they enable the participants in the community of ethical inquiry to raise questions related to the embodied aspect of thinking. In *Harry Stottlemeier's Discovery*, for example, (Lipman and Splitter 1992, 11), the actions of the Aboriginal character Ankuna (an Afro-American called Fran in the original American version) are heavily influenced by both her ethnicity and her liveness, and make little sense if the reader is aware of neither these embodied facts about Ankuna, nor Ankuna's own knowledge of and attitude to these facts about herself.

In dealing with the aspects of thinking separately here, I have possibly given the impression that they are separable in some sort of strong sense. This is not, as I made clear in §3.126, an accurate picture. All thinking involves each of these aspects to some degree, though particular instances of thinking may emphasize one or more aspects over the others. Narrative triggers for discussion have the advantage that, like real lives, they integrate all the aspects, in a way that Kohlbergian dilemmas, maths or physics problems, single teacher questions, conceptual puzzles, descriptions of situations and so on do not. After the presentation of the text, students will ask single questions, but because the questions are rooted in a rich narrative, the discussions will return again and again to this richness, ensuring that the multiple aspects of thinking are visited and revisited.

Narrative texts are able to ensure richness in another way. I have talked of a community of ethical inquiry as a community of inquiry which takes moral and ethical matters as its content. Yet the use of a narrative as the trigger for such an inquiry ensures that moral and ethical matters are not the only ones that get addressed. A narrative, to make any sort of realistic sense, must set moral content in a realistic lifeworld. Many features of this lifeworld, not just the moral ones, may be problematized by the students. The upshot of this is that moral matters are not somewhat artificially separated off from the lived experience of the students, but are discussed in contexts that extend across the students' experiences. The marginalization of moral education that occurs when it is taught as a separate course can be counteracted in this way.

Indeed, the prospect of this marginalization can be further diminished in several ways. If the texts for the community of inquiry mix up ethical with epistemological, metaphysical, logical and aesthetic matters, then moral considerations will be further integrated into the broader lives of the students. The contexts for the narratives can be drawn from the school disciplines, so that ethical inquiry takes place in a scientific or a historical (Sprod 1997b) context. Thus a teacher can attempt to overcome the split between fact learning within a discipline and moral learning in separate (maybe religious education, health or personal and social education) classes, which can be blamed for the failure of many people to think morally within technocratic settings.

One of the important features of the text is that the characters in the narrative ought to be models for the students in the community of inquiry. This is not to say that they need to be perfect, for this would make the text impossibly unrealistic. Hence it would not connect to the lifeworld of the readers, and the mobilizing power of the text would be lost. Nevertheless, the characters can by turns instantiate one or more aspects of reasonableness, so that the community can model their own thinking on it, or fall from reasonableness in some way that is then open to the community to criticize, or haltingly search after reasonableness in a way that enables the community to follow and extend the search (Lipman 1991, 214-219).

In this modelling, the characters are modelling the approach towards communicative autonomy. Characters who are taking charge of their lives through jointly making sense of them with others, characters who are able to form judgements and take actions because they have learned to do so in interaction with each other, are again models for the students in the community of inquiry.

An appropriate text will reflect the real, relatively ordinary, world in its account of the child characters and their setting. In a rich narrative, much of the background is somewhat taken for granted both by the characters within the story and by the reader reading it. Characters move easily against this background, and students will resist the narrative if they cannot, by and large, do so too. The characters will exhibit a narrative unity to their lives, as we usually see them act habitually in a coherent way. None of this is to deny a place to the unusual, the out of character, the surprising from time to time, for the world is like that. It will be these incidents, these realizations, both for the characters themselves, and for the students reading the narrative, that provide the opportunities for the problematization of the taken-for-granted. Within the story, as certain issues are problematized, the characters will investigate them dialogically, in ways indicated above. In this mixture of habituated action and reflection, the stories thus portray the development of children as ethical beings in a

way consistent with the account of virtue ethics modified by discourse ethics developed in Chapter Seven.

I alluded at the start of this section to a tension between the aim of allowing the students to set the agenda for the community of inquiry, and the role of the text in curtailing that choice. By presenting a text to the class that contains certain material, leaving much of it as background and problematizing selected parts, while (necessarily) ignoring everything else, the teacher has restricted the students' ability to set the agenda considerably. Nevertheless, the students are still left with quite a range of subjects about which they can ask questions, and a good degree of latitude as to the specific questions to ask. This is, of course, considerably more latitude than is offered in most classroom activities.

Both my account of how habituated, taken-for-granted knowledge is problematized and made the subject of reflection (§5.232 and §7.11), and my account of learning as a process of connecting new information into an existing construction of the world through social dialogue (§3.23) mandate that the agenda setting stage of the community of inquiry ought to give as much free rein to the students as possible. About the only restriction on questions would come from the necessity of the teacher to enforce respect as a foundational value for the community of inquiry (§6.24). In particular, the teacher ought not to reject questions on the grounds that they are unpromising or irrelevant, for to do so is to model a rejection of respect for the ideas of others. Further, the ideas that lie behind a question may not be obvious to the teacher, but may be explicable by the questioner when dialogue commences, or the community may well be able to find an interesting way of developing the inquiry from the question.

In short, the question agenda must be the students' to set for three reasons: to ensure connection to their lifeworlds; to start inquiry at the edge of their zone of proximal development; and to show respect for the student members of the community.

10.2 The Discussion Phase of the Community of Inquiry

Once the text has been presented and the agenda set, a community of inquiry enters the discussion phase. This phase is said by Lipman to have three stages - solidifying the community (III); using exercises and discussion plans (IV); encouraging further responses (V) - but I think this is somewhat misleading. Unlike the preliminary phase, where reading the text does precede gathering student questions, the three stages of the discussion phase could be used in any order. Indeed, in almost all actual communities of inquiry, these three stages will be intertwined.

The focus in stage III is on group solidarity, with the greatest emphasis laid on the roles of the students. According to Lipman (1991, 242), they will, among other things, articulate disagreements, search for understanding, practice cognitive skills dialogically, join in cooperative reasoning, internalize overt cognitive behaviour, become sensitive to meaningful nuances of contextual differences and follow the argument where it leads. Such a description does not make reference to any leader or shaper of the inquiry, and it is quite possible (albeit not so likely in the classroom) for the processes Lipman lists to occur without any participant in the community taking on the leadership. Certainly, the teacher could join with the students in all these, but this would not require any special role. This is a list of some of the features of communicative action (see §4.122).

On the other hand, though Lipman does not explicitly mention the teacher in listing the procedures in stage IV, the focus is clearly on the teacher's role. The exercises and discussion plans lead to employing questions from the academic tradition, opening students to philosophical alternatives, focusing on specific problems and compelling (Lipman's word) the inquiry to examine the key philosophical concepts (Lipman instances truth, community, personhood, beauty, justice and goodness). These are clearly the roles of a specific person (most likely the teacher) who is coordinating the community of inquiry. They call for certain actions on the part of the teacher, with certain educational goals in mind. The goals may or may not be shared with the other participants: a marker of open strategic action (see §4.121).

Despite the use of the word "encouraging," Stage V mixes teacher and student action. The teacher is to elicit further responses, but the teacher and students together will recognize the synthesis of critical and creative, individual and community, as well as celebrate the deepened sense of meaning. Stage V thus seems to allude both to follow up activities for the community of inquiry, and to an evaluation of the community. As such, it shares some of the features of both stages III and IV.

The tension I have identified between the communicative action of stage III and the open strategic action employed by the teacher in stage IV merely echoes the inherent tension of pedagogic action, as I characterized it in §4.141. Stage III plays the same role of solidifying the community in the classroom microcosm as communicative action does in intersubjectively constructing the lifeworld of society in general. The students achieve the behaviours and aims listed above through engagement with the other students and the teacher in the community of inquiry, but the explicitness with which they approach these aims is quite variable. While they will at times be quite aware of their thinking, and of the skills, tools and purposes which lie behind it, there will be many times when they are "just talking," unaware of the learning going on.

Even the times of awareness (i.e., of metacognition) will usually be at the teacher's behest.

Stage IV underlines the teacher's strategic role in pedagogic action. The teacher joins in the dialogue in quite a different way from the student. While teachers can join in the communicative action, and make the sorts of contributions that any member of the community might make, they must also, simultaneously, be monitoring the conversation from a strategic point of view. The sorts of decisions that are being made as the conversation progresses are aimed at maintaining rigour in the discussion, thus encouraging the moves identified by the teacher as important to this purpose. These moves include process and content oriented moves (though commonly, process and content oriented moves are closely related).

This implies that teachers' engagement in the community of inquiry is distinctly different from that of students. The differences are due to the different social roles played by teachers and students. Teachers are invested with pedagogic authority and the responsibilities that come with it. They are lengthily trained and prepared to exercise pedagogical judgement, just so that they can intervene strategically in the classroom at the appropriate times and in the appropriate way to enhance the induction of their students into society. As we have seen, in the modern West at least, an important part of this induction is the promotion of moral development and communicative autonomy.

This last aim - the development of moral persons with communicative autonomy - implies that the teacher must exercise pedagogical action with an eye to the progress being made by the students. When the students have developed a high degree of communicative autonomy, the teacher can underplay the strategic aspect of pedagogic action and engage in the community more purely on the communicative level. Similarly, the strategic elements of the teacher's action ought gradually to become more transparent, through the students' growing ability to think metacognitively.

This is not to say, however, that the focus of the conversation ought to be split between a metacognitive focus on the mechanics of the discussion and a communicative focus on the content of the discussion. To have such a twin focus would be to destroy the continuity of the substantive inquiry. The procedural elements of the inquiry usually form part of the background against which the inquiry can proceed. Hence the apparently deceptive action of the teacher, who will often (and who ought to) hide the strategic aim of an utterance. To the participants, the asking for a reason for a student's assertion sounds like it is advancing the inquiry (as it is), but the reason will also have been requested as a means of inducing the giving of

reasons for assertions as a matter of course. From time to time, this will need to be made more explicit, but for the most part, such concealed motives are legitimate if they are subsumed into pedagogical action.

Indeed, as Lipman's reference to the teacher's role in employing questions from the philosophical tradition and compelling students to address philosophical themes emphasizes, I have drawn too sharp a distinction between process and content in this account so far. The teacher will not only intervene to provoke better thinking, but also to steer the inquiry into more promising conceptual waters. This is where, in a community of philosophical (including ethical) inquiry at least, the teacher needs to have at least some awareness of, and sensitivity to, the philosophical tradition. The exercises and discussion plans that accompany most *Philosophy for Children* texts assist the teacher in this task.

That the two tasks - procedural and conceptual - are closely related can be shown by considering a few examples. The teacher ought not to ask students to give reasons for all assertions that they make, for some assertions will not be particularly philosophically interesting. So if a student offers the observation that "Essendon are a football team," we would be unlikely to ask for reasons. If the assertion was, rather, "Essendon are a good football team," then asking for reasons would be more philosophically fruitful, as we would be able to enter into a discussion of criteria for judgements of value. Similarly, if a student is relating a tale in which someone lies about an ice cream cone, we are unlikely to judge it worth asking whether distinctions can be drawn between different ice creams (chocolate, strawberry, vanilla), whereas the drawing of distinctions between different sorts of lies (white lies, self-serving lies, lies out of kindness or of malice etc.) would lead to many philosophically intriguing points.

10.21 Following the Inquiry Where it Leads

In the preamble of Chapter Nine, I foreshadowed the need to unpack Lipman's notion of "following the inquiry where it leads" and wondered what the implication of this phrase was for the coverage of curricular material. The phrase might be taken to imply that an inquiry has a purpose of its own, independent of the participants in the inquiry. In some tightly constrained cases, there might be some truth to this. For example, a mathematical inquiry might find that the manipulation of the symbols allows only a certain set of equations to be derived, one after the other. The logical nature of the inquiry compels the inquirers to follow the only path allowed, to a particular conclusion. Now, I am not claiming that this picture is accurate, for it seems to me that even such a tightly constrained inquiry does not lead down a pre-ordained path.

Otherwise, it would be difficult to explain how mathematical inquiries go wrong, and there would be little place for mathematical genius and discovery. But it will serve as a model for an inquiry forcefully leading the inquirers, even though this notion is somewhat problematic.

Let's consider a much less constrained conversation, say around a dinner party table. In most such conversations, it would make sense to say that the discussion leads somewhere, in that it starts with a certain comment, ends with another, and there exists a connection between any one comment and the next, so that we could (sometimes after explanation) see what the connection is. Yet if one participant suddenly says "how did we get to be talking about this?" we will become aware in many cases that the conversation has wandered somewhat aimlessly and contingently about. Despite the fact that segments of the discussion may have made progress towards answering some problem that was raised, there does not seem to have been any overall purpose or direction. Nevertheless, due to the particular inputs made by the participants, the conversation went somewhere, and the participants went there too. In this very weak sense, they followed where the conversation led.

The situation in a classroom community of inquiry ought to lie somewhere in between these two examples, but just where? Who is to be responsible for its direction? In that the community of inquiry starts with a particular question, chosen from the agenda set by the students, we can expect there to be more sense of direction than the dinner party conversation. Yet, as experience in a classroom community of inquiry will soon show, the dialogue there is not usually tightly focused just on answering that question. The inquiry soon throws up more questions, and veers off to address them. It is not uncommon at the end of the inquiry to realize that many of the questions on the agenda have been addressed in some way, and that few of them have been decisively answered. Indeed, students may complain that the discussion has gone round and round in circles. In Sprod (1994a), 20% of 93 Year 7 students made such comments in a survey of likes and dislikes of Philosophy for Children - by far the most commonly cited dislike.⁵

The discussion in a community of inquiry leads where it does under three influences: the impetus given it by the question, the communicative actions of the students, and the pedagogical actions of the teacher. The first of these, of course, has to be interpreted through the second two: it can have no direct input into the discussion.

⁵ Given the training students have had throughout schooling to converge rapidly on right answers, this observation does not imply a weakness in the community of inquiry. Weaning students away from this expectation is an important part of Philosophy for Children. Nevertheless, teachers do need to be aware that some discussions can become somewhat aimless: monitoring the discussion for factors such as this is, of course, an important part of the teacher's pedagogic action.

The inputs of the students are made under two influences: in reaction to the immediately prior inputs of others (like in the dinner party conversation); and under the influence of the question being addressed (less common for the dinner party). It is likely that, for most students, the former influence is stronger than the latter and hence that, all things being equal, the discussion is more likely to meander than to lead. The less experienced the students are in the community of inquiry, the more likely this is to be so.

The teacher, too, is subject to the same two influences, but also to a third: pedagogical concern for the students. Hence, if the inquiry is going to lead anywhere - either to an answer for the question raised; or to a consideration of philosophically important issues implied by the question - the teacher has both a greater responsibility and a greater ability than the students to ensure that it does. Gardner (1995) points out that a failure to realize this means that many teachers, inexperienced in the community of inquiry, preside over a rambling conversation much like my dinner table example. This is not to say that the teacher has to exert tight control over the inquiry, for to do so would subvert many of the pedagogical aims the teacher has in mind. Nor is it to say that the students cannot play an important role here as well. As they become more experienced members of the community, and as the community becomes more self-correcting, they will increasingly do so. But it is to say that the responsibility for oversight of the inquiry does, in the end, remain with the teacher. The inquiry will lead, but only under the watchful eye of the teacher.

Let's look a little more closely at what the pedagogical concern of the teacher entails. In my account of pedagogical action (§4.141), I have analysed the teacher's concern to develop the capacities of students (such as reasonableness and autonomy), but I have not considered pedagogic concern with respect to the content of education. Teachers, particularly in the higher grades, are often constrained somewhat by curricular demands. Yet the community of inquiry, being driven to a large degree by the interests, puzzles and questions of the students, will be, to a certain degree, led in directions that are not able to be pre-specified in curriculum documents.

Does this mean that the teacher ought to exert pedagogical power in order to ensure the community of inquiry covers the ground mapped out by the curriculum? As I stated above, too tight a direction by the teacher will subvert some of the pedagogic aims of the community of inquiry, which rely on the direction, particularly in terms of content, arising from the students. The teacher may remain aware of opportunities to cover curricular ground, and give little nudges in these directions at times, but it does not seem that a community of inquiry could survive with its distinctive strengths intact under a more heavy handed approach. So it appears that there is a *prima facie* clash

between curriculum demands and the community of inquiry as a teaching methodology.

In response to this, I will make two points. Firstly, the community of inquiry is, as I made clear in §9, not the only methodology that ought to be used in a classroom. Some of the other techniques - lecturing, individual research, essay writing, skills practice and so on - may well be better suited to curriculum coverage, when the teacher wants to keep a tight control on the material addressed. Each of these techniques has its advantages, of course, but it is well to be aware that they each also carry certain dangers. In particular, and this is the second point, techniques that are good at covering ground are not necessarily good at promoting real understanding, or motivating the move from understanding to action. This is precisely where the community of inquiry does hold an advantage (§9.21). And this point is, if anything, more important when the subject matter is ethics. It might just be acceptable for students to be able to parrot mathematical, scientific or historical knowledge without deep understanding, but, as Aristotle (1980) puts it, in ethical matters it is not enough for people to “take refuge in theory and think they are being philosophers and will become good in this way” (II.4, 1105b13), especially if the theory is thin and poorly understood.

Thus, an inquiring classroom community will incorporate many teaching methodologies, with the choice of exactly which need to be used when often (but not exclusively) made by the teacher. Yet there is a central place for the community of inquiry, for it is in such a community that students are best able to grapple with ideas and concepts, gaining a greater intersubjective understanding of them. The community need not be entered through the preliminary phase (§10.1), either. The puzzles and questions that lead into a discussion may well arise in the course of using one of the other teaching methodologies. Again, at these junctures, the final decision as to whether to enter the discussion phase of a community of inquiry lies firmly in the hands of the teacher, through pedagogical action.

10.22 Reasonableness through and in the Community of Inquiry

Morally good persons need a well-developed reasonableness (§3.1), for morality requires the ability to come to, and act on, a wide variety of moral judgements and understandings (§5.2). In §6.231, I distinguished presuppositional rationality - which we can, following Habermas, ascribe identically to all normal humans - from situated rationality, which is differentially developed in each human. In education, the question of how presuppositional rationality is transformed into an individual's reasonableness (a preferable term for situated rationality) is of vital importance. Of

course, this process is started long before the child arrives at school, but it continues throughout the years of compulsory education (and beyond, throughout life).

Throughout this thesis, I have been defending the view that Vygotskian learning (outlined in §3.23 and §8.23) is the major basic mechanism by which the development of reasonableness takes place.⁶ Vygotsky, as we saw, claims that such learning leads development. As my account identified five aspects of reasonableness, I will now look in some detail at how the community of inquiry can strengthen each of these aspects, and the ways in which teachers conducting communities of inquiry can enhance this.

In choosing or writing the texts to be used as triggers for the community of inquiry, the five aspects of thinking need to be considered, as I argued in §10.1. There I suggested that a rich narrative, incorporating complex naturalistic situations, opportunities for imaginative extension, varied emotional commitments from the characters and features of the character's embodiment in the world, as well as intellectual puzzles and discussions, was necessary to provoke the full range of reasonableness. Clearly, then, the discussion phase of the community of inquiry will need to build on these features of the text.

10.221 The Critical Aspect of Reasonableness and Transfer

Much attention has been paid in the Philosophy for Children literature to the critical aspect of thinking (see §3.121). Indeed, Lipman's *magnum opus*, *Thinking in Education* (1991) devotes the bulk of its discussion to critical thinking. With regard to critical thinking, Lipman analyses in detail the place of skills, schemata, operations, moves, criteria, megacriteria, metacriteria, logic, standards, reasons, relationships and judgements. I do not wish either to reproduce or to assess this work here, but I shall note a few points.

Lipman, as we saw in §3.121, defines critical thinking as thinking that facilitates judgement because it relies on criteria, is self-correcting, and is sensitive to context. This is a wide and inclusive definition, covering in part the aspect of thinking I label contextual. With its emphasis on judgement, it also looks beyond the narrow skills taught in many versions of critical thinking towards a holistic approach of learning and practicing critical thinking through its naturalistic application within the community of inquiry. As Lipman says:

⁶ This is not to say that there is no place for the sort of interaction with the world that Piaget identifies as important to development, especially in the pre-linguistic months (Kohlberg and Wertsch 1987), but given the focus of this thesis, I shall not pursue this type of learning further.

Instead of directing most of our efforts towards the development of algorithms and heuristics, useful as this may be, we would do better to concentrate on the improvement of reasoning and judgement. The improvement of reasoning involves persistent practice in distinguishing logical from illogical discourse, the acquisition of logical principles, and learning how to apply such principles to actual practice in academic matters and in life generally. The improvement of judgement involves getting students to make and evaluate judgements of practice in a never-ending continuum of judgement, so that the judgements made gradually become more and more reliable and enlightening. (60)

In this passage, there is not a narrow focus on the critical aspect of thinking, as the reference to application in a variety of contexts shows. It is clear that Lipman believes the improvement of critical reasoning is not a matter of isolated, decontextualized skills practice, but of the repeated appropriate use of critical thinking in a variety of contexts. McPeck (1981, 158) criticizes the Critical Thinking movement's approach of teaching thinking as a separate subject, saying that "analysis of critical thinking... makes it abundantly clear that it can only be taught as part of a specific subject and never in isolation." In the community of ethical inquiry, of course, thinking (the critical and the other aspects) is being taught within a specific subject: ethics. As ethics is "the philosophy of morality," this is to follow McPeck's advice about the teaching of both the disciplines and thinking.

One of the most pervasive shortcomings of the way that the traditional disciplines are taught is that they present their material in such a way that its facts and methods are regarded as non-problematic. It is as though the foundation of these disciplines was chiseled out of epistemic bedrock, and all one need learn is what the so-called facts are, and how to use its methods to find out more of them.... A plausible solution to this problem is to make the philosophy of X... an integral part of what it means to "learn X." (1990, 17)

However, it would be a mistake to conclude that all practice of the critical aspect of thinking must be so situated in a specific context that there can be no consideration of its generalization. This is, of course, the mistake that many in the Critical Thinking movement have attributed to McPeck (all reprinted in McPeck 1990; see, for example, Norris 1990; Paul 1990; Siegel 1990). Indeed, the Critical Thinking movement is attacked by McPeck for claiming that the skills of thinking can be taught in a decontextualized form at all.

In [critical thinking's] effort to maximise the number of areas to which its general principles apply, this approach perforce sacrifices genuine effectiveness in all of them. While its prescriptions are generally true, they are also hollow, more truistic than true.... General heuristics are... hollow and virtually useless for specific problems.... Giving people very general principles for solving problems, even with extensive training in them, is like giving people a language with a syntax but no semantic. (1990, 14)

This attack brings to mind the similar Hegelian attack on the emptiness of universalized moral principles I discussed in §7.24, and the approach developed there will also apply here. In both cases, the charge is laid that the more generally

applicable a principle is, the emptier it is. It seems correct to say that working at the highly abstract level makes the problem of application to context-bound particular situations problematic. Yet the thinking that works on specific problems can be generalized and systematized, leading to disciplines with considerable formal power (e.g., formal logic). Just as Benhabib calls for enlarged thinking in relation to general moral principles, I suggest that the answer to the apparent dichotomy between general critical thinking skills and domain-specific thinking skills lies in encouraging, within the community of inquiry, a constant movement between the two.

Research into situated cognition supports this contention. Hennessy (1993), in a review article, avers that situated cognition studies make it

obvious that merely presenting children with new information and experiences... is insufficient to promote learning. To avoid confirmatory bias and to move pupils' primitive thinking forward, new experiences need to be articulated and reasoned about, related to their informal conceptions and generalized to other similar situations. An increasingly large body of research shows that social interaction contributes to children's cognitive development (11)... In sum, learning experiences and activities can be rendered useful and meaningful by the sense made of them by classroom talk (which is generally undervalued as a learning tool). (20)

This explicit articulation and reasoning about new experiences makes the specific skills and abilities used more transparent to the students, and the generalization to similar situations encourages the separation of the abilities from a narrowly specific context. The call for social interaction and classroom talk points towards a pedagogy like the community of inquiry. Later, Hennessy gives more details of the type of classroom talk and social interaction the studies she is surveying recommend.

A critical insight derived from the situated cognition research is that problems emerge out of dilemmas and learning arises when means are sought to resolve those dilemmas. The implications are that formal educational settings need to encourage active intellectual engagement in... thinking... We should encourage schoolchildren to formulate, attempt to solve and communicate their discoveries about questions arising in their classrooms, playgrounds and homes... Problems should be ones pupils want to solve, which are real and relevant to them, which engage their interest, and for which they can take responsibility. (33)

With its emphasis on children formulating the problems, we are reminded of the preliminary stages of the community of inquiry, discussed in §10.1. Although these problems are drawn from fictional narratives rather than the lives of children (as Hennessy recommends), the use of anecdotes within the inquiry helps to bridge this gap. I will discuss this point further in §10.223. But while the teacher ought not to choose the problems and puzzles to be discussed, Hennessy sees a role for the teacher in the encouragement and scaffolding of the inquiry. Rogoff and Gardner (1984, 97-98) highlight the importance of the joint construction of the conditions for

generalization or transfer by the teacher and learner, in a way which supports my analysis of pedagogic action in §4.141:

Inherent in instruction is the construction of a context in which the new information is made compatible with the learner's current knowledge and skills... [by means of which] the teacher is guiding the learner in the generalization to the new problem.

It is the ability of the community of inquiry to focus in turn on the attempt to come to an understanding or a solution to a substantive problem, and then to focus metacognitively on the tools of inquiry which have been used in that attempt which provides the bridge between domain-specific engagement in inquiry and domain-general transfer of the means of inquiry. The clear implication for teachers conducting a community of inquiry is that they ought to be constantly aware of opportunities to nudge the level of inquiry from the specific and relatively unreflective use of the critical aspect of reasoning to explicit metacognitive inquiry into the tools of reasoning and back again at appropriate times. The appropriate time is, of course, a matter for nuanced and situated judgement, and must be based on the students' contributions. The time to move to the metacognitive is when the use of tools is getting muddled, and the time to reapply the tools is when the metacognitive reflection is getting too disconnected from real examples. In the Philosophy for Children support materials (the manuals), many of the exercises are designed to assist teachers to make the moves in either direction. Perkins and Salomon (1989) claim that this sort of reflective addressing of the application of cognitive tools to other situations does enhance the possibility of transfer.

In this discussion of transfer, I have concentrated mostly on the critical aspect of reasonableness, partly because this is also the focus of the authors discussed. But McPeck's critique, and the conclusions reached above concerning it, are applicable to all five aspects of reasonableness, especially as they are intertwined in any actual example of thinking. It is to this intertwining and the implications for the leader of the community of inquiry that I now turn.

10.222 Admixture of the Aspects of Reasonableness in the Discussion

Lipman, who defines creative thinking "as thinking conducive to judgement, guided by context, self-transcending, and sensitive to criteria" (193 - see §3.122), recognizes this intertwining with respect to the two aspects that form the focus of his book.

There is no creative thinking that is not shot through with critical judgements, just as there is no critical judgement that is not shot through with creative judgements. We can, of course, construct abstract ideal types in which pure forms of thinking are delineated, but in actuality admixture is the rule. (194)

Likewise, I claim that thinking must be “shot through” with the committed aspect (without which there would be no motivation to think at all), the contextual aspect (as Lipman recognizes in his definitions of both critical and creative thinking), and the embodied aspect (which recognizes that it is *we*, as body-and-mind persons, who are situated within the context). Each of these aspects may come to the forefront of our thinking, depending on the details of our inquiry, and each may be turned to, reflexively, to become the “abstract, ideal type” that we consider for the moment, before it returns to the background as the inquiry progresses. I have discussed the theoretical underpinning for this claim in §3.126.

Since the rich narrative of the texts chosen will also contain the same admixture, students will most likely raise questions which have the potential to access each of the aspects of reasonableness at different times. This will inevitably lead to a rich mix in the discussion that follows. However, different questions will highlight different aspects to varying degrees, and it should be one of the aims of the teacher to build on this by similarly foregrounding each of the aspects at appropriate times, as well as facilitating inquiry into their interactions.

Transcripts of discussions within a community of inquiry can be analysed in various ways. As presented in §8.23, I have outlined the structure of such discussions in previous work (Sprod 1994c; 1997d; 1998). They can be divided into segments, containing from a few to several dozen utterances, which I called “epistemic episodes.” These I defined as “those distinct segments of a discussion that approach the pursuit of knowledge in a particular way” (1994, 2). While the focus in that work was mainly on the critical aspect of thinking and the dynamics of the discussion, I did look at the way in which the cognitive and affective aspects of the discussion intertwined.

This empirical research into the structure and development of a community of inquiry within a Year 7 science classroom was completed well before I carried out the conceptual analysis of reasonableness presented in §3.1. Hence, the analytic categories I used in that study do not match those used here. The major distinction that I drew in that study was between cognitive and affective thinking. Roughly speaking, I would now further categorize thinking that I called cognitive in that study as displaying mainly either the critical, the creative or the contextual aspect of thinking, while episodes of affective thinking would most likely now be categorized as mainly displaying the committed aspect of thinking. Within the cognitive category, I made a distinction by noting whether the cognition was largely turned towards the process of the thinking or the substantive content. The latter are good candidates for classification as foregrounding the contextual aspect of thinking. With the benefit of

the more careful conceptual analysis I have carried out here, I would probably also have identified some episodes that displayed the embodied aspect of thinking.

Leaving aside these analytical differences, there was clear evidence that different epistemic episodes tended to focus now more on the cognitive element (substantive or procedural), now more on the affective, and that some episodes combined both (Sprod 1994c, Table 5.1). The close analysis of one discussion which is presented in that table illustrates this point. The discussion had been preceded by an experiment using a van der Graaff generator, and the broad content categories I used covered static electricity, current electricity, energy conservation and the mechanics of experiments. The process categories concerned testing, appeals to experience, reasons and deduction. Affective issues were not differentiated.

Of the fifty seven epistemic episodes identified in the discussion, twenty eight were placed in one of these categories only. This does not imply that they contained only one element of thinking, but rather that the one element was clearly to the fore. For example, episode 20 (substantive) is concerned with what static electricity can do, episode 28 (procedural) addresses what a fair test is, episode 4 (classified affective, but I would now call it embodied) reports what it felt like to be charged by the generator. The other twenty nine episodes sufficiently displayed a combination of different elements of thinking that they were classified under two or more elements. In episode 12, for example, the class is trying to ascertain whether a generalization they have just made will apply to a particular unusual case, and thus they mix the contextualized aspect of thinking (about the specifics of the generalization and the application) with explicitly considered reasoning (procedural element = the critical aspect). In episode 29, a student explores the connection between the implications of an experiment (procedural element = critical aspect) and the harm it may do the student involved (affective element = committed aspect).

This empirical research highlights the fact that the extent to which each of the five aspects of reasonableness dominates a particular socially constructed episode of overt thinking varies quite markedly. Sometimes, one aspect is clearly predominant; at other times, several of them are very closely combined. Yet, one of the marks of the epistemic episode is that it maintains, over several (maybe many) utterances from different participants in the dialogue, a consistency of aspectual mix. In engaging in a specific epistemic strategy, we engage the appropriate aspect(s) of reasonableness. As the strategy changes and the dialogue shifts into a new epistemic episode, so too may the aspects of reasonableness that are being wielded. Awareness of this epistemic episodic structure of discussions, together with attention to the features that are being picked out in each episode, would enable teachers to focus their comments within an

episode on sustaining the aspectual nature of the inquiry within the episode, and to link it reflectively to other episodes. In this way, all the five aspects of reasonableness can be strengthened, students learn how to engage them appropriately in context, and the progress of the inquiry will be served.

Of course, it is not necessarily true that all the features of any classroom discussion will serve inquiry and meet the pedagogical aims of the teacher. Clearly, teachers need to be vigilant to steer the discussion away from areas that have negative effects, such as loose free association of ideas, so that the students' comments do not connect to each other, or "'tis!," "'tisn't!" exchanges that do not engage in any reasoned argument. Yet to make such a judgement requires that teachers have a clear idea about what the threats to worthwhile dialogue are. I will now turn my attention to two threats that have been identified, and assess them in the light of my aspectual account of reasonableness.

10.223 The Place of Anecdotes and Psychologizing in the Inquiry

The distinction between a discussion within a community of inquiry which is inquiring properly and one which isn't has been marked in a number of different ways. Splitter and Sharp (1995, 34-40) mark it by calling the former "dialogue" and the latter "conversation" (see also Gardner 1995), though it is worth noting that some other authors (e.g., MacColl 1992; Reed 1992) have used "conversation" in a sense closer to the former. Without entering into this debate, I shall here use "dialogue" and "conversation" as a shorthand to mark this distinction. For Splitter and Sharp (op. cit., 34), the four "necessary and, perhaps, jointly sufficient" characteristics of a dialogue, are: a focus on a problematic or contestable topic; self-correction; an egalitarian structure; and guidance of the inquiry by the mutual interests of its members.

In this section I want to concentrate on the first characteristic: the question of how teachers are to determine whether the discussion is sufficiently focused to count as a dialogue. Splitter and Sharp explicitly say that their first condition ought to be interpreted liberally, but they obviously wish to exclude features of a conversation that mean that it lacks the sort of focus they have identified. Philosophy for Children theorists commonly claim that the focus must be philosophical, and that this means teachers ought to discourage non-philosophical discussion in their community of inquiry. Two of the threats to the philosophical nature of the inquiry that have been identified are the telling of anecdotes and "psychologizing" (I shall document these claims and the reasons advanced for them below). Broadly speaking, I agree that a community of inquiry needs a philosophical focus, but I shall advance arguments that

these exclusions arise largely from a relatively narrow account of reasonableness, and that my broader characterization of reasonableness warrants their inclusion, provided certain precautions are observed. After considering the arguments for each in turn, I will turn to the claim that anecdotes and psychologizing are of special importance if the focus of the inquiry is on ethical matters.

Students will commonly respond to an incident in the trigger text by telling an anecdote from their own experience. Other students will then want to tell their own anecdotes, in a sort of free association. The inquiry has the potential to degenerate into a string of unconnected anecdotes; a mere conversation. For this reason, the telling of anecdotes has been frowned upon in the community of inquiry, unless the anecdote can be tightly tied to the inquiry as an example or counterexample of the point at hand. Here is a cautionary note from Reed (1992, 154), concerning:

the overuse of personal anecdotes. When we inquire together, it is almost impossible not to argue from personal experience, to relate bits and pieces of our histories. There is nothing wrong with using anecdotes to facilitate the process of inquiry. The problem, however, is that all too frequently inquiry tends to degenerate into a mere recitation of personal history. A good rule of thumb might be to eliminate all anecdotal reports unless there is a compelling reason to share them.

Elsewhere, Reed (1992) has argued for the place of anecdotes as a means of producing a realignment of the dialogue, directing it into new avenues of exploration, but this again depends on the anecdote contributing directly to a fruitful direction for inquiry; Reed offers criteria by which it can be judged as such by the teacher.

The thrust of these arguments against permitting the telling of anecdotes is that they interfere with the “process of inquiry,” where inquiry is conceived narrowly as involving only the critical and creative aspects of thinking. By including consideration of the contextualized, committed and embodied aspects of reasonableness, I maintain that there is a larger role for the “recitation of personal history.”

Anecdotes are episodes from the lived histories of the students. They are thus rich in contextual detail. This benefits those who retell them, who through the telling connect the substance of the inquiry to their own experience and situatedness in the world. Anecdotes are also rich in emotional attachments. They are rarely mere recitations of the facts, usually providing a way for those telling them to work through and even, often enough, discover their emotional reactions to the events in their lives. They are motivational for the tellers, who discover in the telling of the anecdote that they are now personally connected to the discussion. In anecdotes, students bring *themselves*, as whole, embodied persons, to the inquiry, which becomes woven into their life’s narrative.

To concentrate on the tellers of the anecdotes is to miss the greater part of their benefit. The telling of anecdotes also benefits the listeners, for they gain an insight into the world inhabited by others, which has the effect of changing their own worlds. The sharing of anecdotes is a part of the process of creating an intersubjective world, in which it is possible to understand others. In hearing of the way you see your surroundings, feel about the events in your life and react to them, I am not merely learning about you, but also envisaging for myself alternative possibilities for my own dwelling in the world.

Anecdotes, even a series of somewhat loosely connected ones, can serve all these purposes, though it must be emphasized that the teacher needs to exercise pedagogical judgement as to whether the gains outweigh the disadvantages pointed out above. Clearly, an anecdote performs these services best when it meshes well into the inquiry and contributes to its further advancement, for these are the conditions under which the whole community is most likely to remain engaged and to explore the implications of the anecdotal narratives. These considerations speak against the allowing of “free associating” from one anecdote to another, for an open slather attitude to anecdotes has dangers of disconnection from the dynamics of the inquiry. Nevertheless, the “elimination of all anecdotal reports” has its own dangers, leading to the over-formalization of the inquiry in terms of the critical (and perhaps creative) aspects of thinking, thus cutting the discourse free from the lives and concerns of the participants. One of the capacities that students have to learn is that of using anecdotes in a way that does not just connect their experience to the conversation, but also helps to advance the inquiry itself. In pursuit of this cause, teachers can ask students who are offering anecdotes to make clear the ways in which their anecdote sheds light on the topic in hand, or request that other students look for the connection.

The second danger that teachers have been warned to avoid is “psychologizing.” This phrase refers to allowing students to engage in amateur analysis, turning the community of inquiry into “encounter-like discussions which probe the personal details of one another’s lives” (Splitter and Sharp 1995, 181). The warning is wise, for to engage in an intrusive exploration of students’ deepest feelings in their presence without proper training does risk doing damage. Yet, while Splitter and Sharp see that “it might be *philosophically* appropriate to examine the thoughts, character and actions of a specific fictional character... [knowing] that such examination cannot be harmful or invasive” (185n40, *italics added*), there is a resistance to allowing a folk *psychological* analysis of the characters, on the grounds that it is inimical to inquiry.

It is worth noting that psychologizing can be seen as an extension of the sharing of life experiences, be they anecdotes or fictional vignettes, in that the students in the inquiry

take up the anecdotal narrative and explore the psychological factors that lie behind it. Thus, in that psychologizing leads to an investigation into the events retold, it helps to induct the anecdote into the inquiry. As the benefits of anecdotes that I discussed above depend to some degree on the exploration of the anecdote within the inquiry, then allowing a psychological approach will enhance these benefits.

There are two reasons commonly advanced why psychologizing is a bad thing. The first, mentioned by Splitter and Sharp, is that there are dangers of psychological damage to individuals within the community if the focus is turned on them. I certainly endorse the warning to be careful in these situations, and the advice that such probing should rather be done on fictional characters, so as to avoid this danger, is apposite. However, there is a distinction that can be drawn between everyday psychology and abnormal psychology. If the probing is into common psychological factors that are likely to be felt by most of the community, then psychologizing contributes to the building of an intersubjective knowledge of persons and their position in the world. It builds community. The dangers are much greater when the factors under discussion are seen by many in the community as unusual or “weird.” It is then that teachers must exercise caution and care for the members of the community.

The previous paragraph reads largely like a “tip for teachers,” but my second point has more philosophical bite. The disapproval of psychologizing, whether it be based on anecdotes from the community or vignettes from the text, also arises from a feeling that the inquiry has been sidetracked and dissipated, forced away from philosophically interesting and worthwhile matters. I argue that this fear is based on the impoverished view of rationality as critical and creative thinking, and that my concept of reasonableness gives us philosophical reasons for incorporating psychologizing into the inquiry, as one of many approaches. As the first step, I will present an example of what an episode of psychologizing might look like, and the common arguments against allowing it.

Students often ask questions about motivational factors that lie behind the actions of fictional characters. The reason for the reluctance to allow such inquiry I have heard given at Philosophy for Children training courses is that it leads to a barren listing of possible psychological motives, with no way to decide between the alternatives. The dialogue degenerates into a series of “maybes.” Because such psychological questions are empirical and not conceptual, it is claimed, in the absence of textual clues, questions about motivation remain undecidable.

The hidden premise of this argument is this: rationality is primarily about critical and creative thinking. In generating possibilities, students are admittedly exercising

creative thinking, but it is only in conceptual matters that possibilities can be critically discussed; in empirical matters, hard empirical evidence is needed for progress to be made in critical discussion. Therefore, in the absence of such evidence, such discussions should be avoided. In the light of a richer account of reasonableness, which recognizes the committed, contextual and embodied aspects, I argue that such discussions still have value for the community of inquiry.

Although students have neither the empirical evidence, nor the psychological expertise to definitively answer the sorts of psychological questions I am referring to here, they still need to become capable folk psychologists in order to operate in the lifeworld. This entails building up their intersubjective knowledge of both themselves and others as agents within the world. This is especially true in the case of ethical (or practical) questions, as they will have to act in social situations where understanding both themselves and others will be of vital importance. They need thicken their understanding of how emotional, contextual and bodily factors frame situations. They need practice at imagining the situations and motivations of others (utilizing the creative, committed and contextual aspects of thinking), and of gaining greater insight into their own motivations (where the embodied aspect of their own thinking will also enter).

Of course, ordinary immersion in the lifeworld makes children reasonably competent folk psychologists at quite a young age, but reflective engagement in psychological inquiry, especially when mixed with philosophical inquiry into the concepts of ethical personhood, will enable them to hone these lifeworld understandings. Indeed, given my richer conception of reasonableness, there are many links between philosophical and psychological inquiry. For example, speculative answers about psychological motivations may not be decidable, but are certainly criticizable on multiple criteria, such as likeliness, consistency with known character traits and so on. Finding and evaluating such criteria provides rich philosophical possibilities. More importantly, our key ethical concepts are not, as Kant and Kohlberg would have it, context-free principles of Reason, but rather reflectively constructed from paradigm scenarios, as described by Matthews and Johnson (see §5.1 and §6.2322 respectively). These scenarios are based on contextually situated, emotionally charged, embodied experiences, initially from our own lives but greatly enriched by consideration of, and engagement with, similar experiences in the lives of real or fictional others. The construction of ethical concepts thus requires some facility in both philosophical and psychological inquiry.

This leads me to the claim that there are special reasons why anecdotes and psychologizing ought not to be excluded from a community of *ethical* inquiry. More

than many of the other subdisciplines of philosophy, ethics is centrally concerned with our relationships with others. While it might be claimed with some justice that logic, for example, draws most heavily on the critical aspect of reasonableness, my virtue ethics/discourse ethics account (Chapter Seven) shows that thinking about ethics has much to do with contexts, commitments and embodiment, especially in interrelationships with others. As I have shown, this provides a space for anecdotes and psychological inquiry.

Glaser (1998) discusses the sort of imaginative projection that ethical thinking requires, referring to Arendt's "visiting imagination" and Nussbaum's "judicious spectatorship" (which Nussbaum in turn derives from Adam Smith's work). Both these concepts attempt to capture the sense in which we can make judgements about the experience and actions of another in ethical situations. It is not enough for me to put myself imaginatively in the shoes of the other, because as myself, I will not construe the situation as the other does. I will end up making a monological judgement. Nor can I imaginatively project myself onto the other, taking on all their views and beliefs. If I succeed, then the judgement I make is theirs, not mine, and again monological. In either case, the judgement is likely to be inappropriately biased through having the emotions of a participant, rather than a spectator who can "omit that portion of the emotion that derives from our personal interest in our own well-being" (Nussbaum 1995, 74).

Thus, I need to imaginatively place myself alongside the other as a friendly and sympathetic companion, knowledgeable about their views but retaining my own judgement and a certain detachment from their predicament. As Smith says: "The compassion of the spectator must arise altogether from the consideration of what he himself would feel if he were reduced to the same unhappy situation, and, what is perhaps impossible, was at the same time able to regard it with his present reason and judgment" (quoted by Nussbaum 1995, 73-74). To have this knowledge requires that I can "imaginatively construct the multiple ways a situation may present in order that [I] may stand alongside them" (ibid, 20). It is this that can be practiced in the community of inquiry, where engagement with others in dialogue gives us the capability to gain an insight and understanding of others, whilst simultaneously strengthening our own ability to make judgements that take account of these understandings. Both anecdotes and psychologizing contribute to the insight and understanding, and incorporating them into the inquiry aids the development of judgement.

In the engagement with literature (Nussbaum 1995) such as the texts that used as triggers for the community, these inquiries are opportunities for sharpening moral

discernment without potentially harmful consequences to real persons. As is not the case for solitary reading, however, ideas can be tested in the forge of intersubjective discussion and enriched by the input of others (Glaser 1998, 21). Nussbaum recognizes this, saying that “ideally the process of reading must be completed by a conversation among readers” (op. cit., 75). Indeed, in hearing and grappling with the anecdotes of others, a better grasp of the psychological reactions of real others, outside the fictional world, to situations can be gained (provided those others are protected from harm).

In the foregoing, I have suggested that discussions within the community of inquiry may justifiably be widened beyond the scope normally allotted to them, and that this is especially so in an ethical inquiry. My account of reasonableness draws attention to the committed, contextual and embodied aspects of thinking. While these aspects have a role to play in all thinking, their role is more important when the focus of that thinking is not merely what to believe, but also what to do. Therefore, in a community of ethical inquiry, those factors which mobilize commitment, tie the discussion more closely to real and imagined contexts, bring participants in as real embodied actors and connect the discussion to the motivations and experiences of action in the world will enhance the dialogue and increase the probability that its conclusions are carried through into the lived lives of the students. Thus they can develop not only reasonableness in thinking, but also reasonableness in action.

10.23 Moral Correctness and the Community of Ethical Inquiry

The major focus of the community of ethical inquiry, of course, is the development of ethical persons. Reasonableness is, as we have seen, an essential part of this, but there is a sense in which reasonableness is more process oriented, and ethical persons need also to have substantive moral views and beliefs. Hence I now turn to the question of how the community of ethical inquiry can lead to students acquiring adequate moral knowledge and beliefs.

This, of course, raises the vexed question of how we can recognize moral knowledge as adequate, particularly in a pluralist society. For some, moral knowledge will only be adequate if it is correct; that is, if it corresponds to moral truth. The problem with this, of course, is determining which moral views are in fact true.

In the course of this thesis (§6.24 and §7.21), I have sought to establish the position that an insistence on, and hence an inculcation of, procedural respect is not only justified, but also essential to the educative endeavour. This is especially true of the community of ethical inquiry. Such an insistence does not amount to indoctrination, partly because the substantive notions of respect that underlie the procedures are

potentially open to inquiry within the community, and this can lead to both discussion of, and modifications to, the procedures of the community (§9.413).

Indeed, the inculcation of the procedures and values of respect for others within the community of inquiry is just a special case of the inculcation of values within the lifeworld. Lifeworld values can vary on several dimensions: their range of acceptance (they can be relatively widely accepted throughout the world, accepted only within a given society or particular to some smaller grouping such as a subculture or even a family); their specificity (they can apply to all persons, or only to particular persons); and their strength (from strict taboo to mild disapproval). What they all have in common is that they are taken-for-granted within their range.

Lifeworld values in general can emerge from the background and become problematized at any time. In the community of ethical inquiry, this is more likely to happen to both the inculcated values and lifeworld values brought into the community by the participants, simply because the asking of questions about the problematic is so encouraged. To concentrate for a moment on the procedural values of the community, we can comment that the problematizing of a particular value, say, that one ought to listen when another is speaking, can take one of two directions, identified by Habermas in the distinction of discourses of justification from discourses of application (§7.24). In the former, the community will assess whether the overall value is a worthy one: is it ethically required to listen when another is speaking? In justification, we seek to redeem the general value. But even if we agree that, other things being equal, listening to a speaker is morally better than ignoring them, this does not decide the details of the application of the rule. Does it require that all present look at the speaker? On what grounds, if any, may another interrupt the speaker? Even if all present agree that one ought to listen to a speaker, there are grounds for much disagreement on the answers to questions such as these.

The implication of this for the teacher leading a community of ethical inquiry is that all questions concerning norms - be they procedural or substantive - ought to be grist for the mill of the inquiry. Unlike in most classroom activities, the child who asks "Why are we doing this?" is not leading the class off on a wild goose chase (even if that might be the intention). If the question does raise a matter of value, and if the community is interested in following the inquiry in that direction, then the question is not only legitimate, but also quite central to the business of ethical inquiry. Indeed, even if the rest of the class is somewhat dismissive of the question, then the teacher might legitimately ask for one of them to explain to the questioner why the values that lie behind the question are justified. If they can easily do so, then the matter may be

passed over. It may, however, transpire that the matter is not as straightforward as the answerer assumes, and the ensuing discussion can easily be a fruitful one.

This leads us into consideration of the question of tolerance: how we are to decide which views ought to be tolerated within the community of inquiry. That such a community ought to tolerate a considerable degree of difference of opinion is clear, for intolerance is inimical to the ethos of open inquiry. Nevertheless, there must be some limits to tolerance. A first answer to the question is to draw a distinction between lazy tolerance and critical tolerance. We tolerate a view lazily if we merely accept the right of someone else to express it. We show critical tolerance if we accept the right of someone else to express a view provided they are willing if required to defend the view (equally, we must be willing to consider seriously the reasons advanced in that defence). In Habermasian terms, we tolerate views where the holder of the view will honour this speech act immanent obligation (§4.124). The obligation also extends to a willingness (on both sides) to change the view if the reasons for doing so are sufficiently compelling.

But this first answer is not enough, for the holder of a view might be willing and able to defend it, yet the view itself might be “beyond the pale.” We need an additional criterion for separating views that ought to be tolerated, even if they are bizarrely different, from those that ought not. Here, I turn again to the foundational value - respect (§6.24). We must tolerate a view which does not violate respect for others, and which is defended respectfully. The teacher (as well as any other student) has the right to demand that non-respectful views, or non-respectful defences of them, are withdrawn from the community, or recast in respectful terms.

Students within the community of ethical inquiry can, though a consideration of the tractability of certain issues to approaches to consensus, strengthen their feel for the difference between what I have dubbed (in §7.24) rightish and goodish issues. As we saw there, some issues are more liable to result in consensus (the rightish), whereas others are prone to lead to agreements to disagree. Issues of the good life, they are liable to decide, call for a greater degree of tolerance than issues of justice. Yet, because there is no “sharp cut” to be made between the right and the good, even this conclusion can be contentious. In any case, a narrowly based consensus may arise merely because the students’ views are embedded in similar background assumptions and lifeworlds.

Let’s leave aside cases where the community comes to a decision to tolerate a difference of views respectfully, and look further at the case where the community of ethical inquiry has come to consensus on a matter either of justification or of

application of a value (§6.21). There is still the question as to the weight that this agreement carries with respect to normative correctness (§6.22 and §7.23). A classroom community of ethical inquiry is an infinitesimally small discourse when compared to the ideal practical discourse that Habermas claims as the only warrant of normative correctness. Just because a class of (say) thirty 12 year old white middle class Australian students agrees after 30 minutes of discussion that the habitat of the bilbies ought not to be destroyed, does this provide any warrant for members of that class believing in the correctness of this norm?

Clearly, whether there is a warrant or not, in the absence of any other input a student who does assent in discussion to this norm will as a matter of fact believe it. For us, the question is not about whether we are right to hold the values that we have come to hold (for we do in fact hold them), although others may question this. Rather, we need to be concerned about the manner in which we hold them. If the community of ethical inquiry is working well, one of the values that will be learned (possibly implicitly, but always with the potential for explicit discussion and endorsement) is that, while we must adhere to, and act on, our values, we ought also to be open to the possibility that further discussion might lead us to change our minds, provided the reasons for doing so are sufficiently compelling.

Nevertheless, teachers ought to take seriously the homogeneity of such a classroom community of inquiry, and the potential that its homogeneity has for leading to a narrow, poorly considered moral consensus. For, even if the process of inquiry within the class is exemplary, in the absence of a sufficient degree of plurality of views, such a narrow consensus is the most likely outcome. In §7.23, I discussed the notions of progress in virtue ethics. Included in that discussion was a consideration of the “dark side of community” (Noddings 1996), with its potential for narrowness and oppressive conformity. This is precisely the situation that the classroom community of ethical inquiry can face. If the children are drawn from a particular social grouping, their deliberations are likely to consider only the lifeworld norms of that group. In the absence of diversity within the classroom community, this can lead to the easy acceptance of norms that would be much more problematic if discordant views had been introduced to the discussion.

I have argued that progress in virtue ethics requires wider and wider consideration of views. The same applies to the classroom community of ethical inquiry: if we are to have any hope that the deliberations of the community are leading to progress in the students’ moral outlooks, then it must be based on an increasingly wide consideration of alternative ideas. A classroom drawn from a narrow social grouping can instantiate many of the limitations that Benhabib (1992, 141) identifies in the isolated thinker:

A major mistake of Kantian moral theory is to assume that the principles of enlarged thought can be realised via the isolated thought experiments of a thinker. These solitary thought experiments often substitute the standpoint of one privileged part for that of the whole.... we have to see that "to think from the standpoint of everyone else" entails sharing a public culture such that everyone else can articulate indeed what they think and what their perspectives are.

If we are to encourage an "enlarged mentality" (a concept Benhabib, *op. cit.*, 122-123, borrows from Arendt), then we need views injected into the community that might arise from beyond "one privileged part" of one society. Such an enlarged mentality (§6.23) is the road to a more universal perspective, and hence to a greater hope of moral correctness (Benhabib, *op. cit.*, 136-137). But if the classroom community is not itself taking in new members, from a diversity of backgrounds, from where will this width come? There are two answers to this question.

The first highlights again the role of the trigger text in the community. Fictional characters can express, or report, dissident views and thus introduce diversity into the community. However, there are two problems with this solution, and both problems limit the extent to which the text can overcome this difficulty. Firstly, the writer of the text, separated in space and time from the communities that will use the text, cannot easily anticipate the lacunae in the community's knowledge. Secondly, if the text is to retain the power of narrative, it must read reasonably naturally. It must represent a recognizable fragment of the lifeworld, and that lifeworld must be sufficiently familiar to the readers to enlist them. With the need to introduce philosophical considerations a central concern of the writer, there are already dangers of disconnection from the readership, who are liable to criticize the texts for being "all about schoolwork,... not natural, doesn't relate to our generation, uninteresting" (quotes from a survey reported in Sprod 1994a, 61, where 17% of Year 7s surveyed criticized the text in similar terms). Authors must be wary that the injection of too many divergent views into a text might destroy its aesthetic and narrative integrity.

The second, and to my mind better, answer focuses on the teacher. Phillips (1994), in an article entitled *A sincere word for the devil's advocate*, points out that the community itself has interests that are not identical to the interests of the individuals in it. One of these interests, he says, is its "epistemic interest," which is served by ensuring that possible objections to the conclusions reached have been adequately considered. No member of the community has an interest in advancing positions opposed to their own position, for sincerity is an important attribute for individuals.⁷

⁷ This is too simplistic in several ways, and I will make a couple of observations. Firstly, this assumes that the interests and positions of the members of the community are fixed. In fact, these interests and positions, as well as wants, needs and so on, can themselves be modified through participation in the community of inquiry (§7.11). Secondly, as Glaser points out (pers. comm.), there is a sense in which every member of the community has an interest in considering possible

Nevertheless, it is in the interests of the community as a whole that ideas are fully tested before gaining assent. Phillips says:

We want a broad and generous devil's advocate who is not merely stung by her community's too-hasty acceptance of this or that position and who responds by raising objections to the popular view; we want one who responds by undertaking to lead the community into exploring alternative views. (18)

I want to be clear about how I understand the notion of “a broad and generous devil's advocate.” The phrase, as Phillips notes, derives from the Catholic Church: the devil's advocate (properly called the *promoter fidei*) is assigned the role of advancing all the reasons why a proposed saint ought not to be canonized. In more general usage, the phrase has come to denote someone who argues a case they do not hold. In my interpretation, it does not matter whether the view is argued hypothetically, or on behalf of a particular concrete other who is not present, or somewhat deceptively, by someone who pretends that they do hold a view that they do not in fact hold. Glaser (1998), whose views I will consider below, restricts the phrase to the last case only, but I think this is mistaken. Few would believe that the *promoter fidei* is personally committed to all the objections he raises, even though they are presented as if he did. The hypothetical nature of the role is no secret, so no deception is practiced.

Generally, the only member of a classroom community of inquiry who can be counted on to come to it with a commitment to the interests of the community as a whole is the teacher. This is part of the pedagogical role. Students come into the community representing themselves, whereas the teacher has a mandated role as the representative of the community of inquiry itself. Any student may choose to take on the “devil's advocacy” if they wish, and it is certainly true that one of the aims of the community of inquiry that students do learn to take into account possible objections to their view. The teacher, however, has a duty to raise objections once it is obvious that none of the students are going to introduce material that throws doubt on the conclusions being reached. This will be true even if that conclusion is the one held by the teacher. Teachers, of course, need not, and generally ought not, to pretend that the objections are their own, but this does not mean that they cannot if there are good strategic reasons for doing so.

As I indicated, one of the dispositions we would want students to learn is the somewhat difficult and counterintuitive one of overcoming confirmation bias by searching out and seriously considering objections to their views. Students will seldom be either disposed or able to do this when they enter the community of inquiry. In view of the Vygotskian account of learning that underpins this thesis (§3.23), this

objections to their position, if their aim is to find the best position, rather than to merely defend their present position. I shall return to this latter point in due course.

will be learned through participating in a community which does consider objections. This benefit accrues even if the person playing the devil's advocate does so deceptively, but other considerations speak in favour of a hypothetical or a representative role. The first has to do with trust.

Glaser (1998, 316-317) analyses Phillips' article in terms of trust. She claims that no member of the community, including the teacher, can act in a way that violates the "stronger condition of trust that underpins dialogical action," and that playing the devil's advocate does betray trust because the devil's advocate argues a case that is not sincerely held. I have noted above that Glaser has a narrow interpretation of the role of the devil's advocate, but even if we for the moment accept her version, does this mean that the teacher can never advance arguments for a view they do not hold, as if they did hold it? In her analysis, I maintain that Glaser collapses the role of the teacher into that of any member of the community. She denies the pedagogical authority of the teacher: a role that entails, as I argued in §4.14, a different variety of trust from that which Glaser identifies. Her trust is the trust extended to a fellow member of the community - that they will be sincere - which ought to apply to the student members of the community. The teacher, on the other hand, is mandated to take open strategic action, and the trust that students extend to the teacher is the belief that the teacher's actions will be educational. The teacher has a duty to meet that trust.

Yet, while it may be occasionally acceptable for the teacher to pretend to hold a view, so as to model diversity, other considerations speak against doing so. If the teacher is discovered arguing deceptively, it will break down the trust that Glaser identifies, and model "bush lawyerism" (§9.22). Further, such deceptive argumentation only allows an implicit Vygotskian inculcation of the consideration of objections. If we want students to obtain a more transparent and explicit understanding of the importance of taking possible counter-arguments into account, then it must be clear that the teacher is doing so despite not agreeing with them. This is only achieved if the teacher makes it clear that the objections are hypothetical, or derived from a concrete other who is not present to represent their own view. Generally, we can conclude, the devil's advocate ought not to be a deceptive devil's advocate. Nevertheless, there is an important role for the teacher as devil's advocate.

Clearly, the duty to act as the devil's advocate does not override all other duties to the community. The teacher, as leader of the community, needs to exercise pedagogic action in balancing the competing interests of the community. For example, the discussion may have been going on for some time, and the introduction of yet another objection may pose the danger of making it too tedious, or the dynamics of the dialogue may have pushed it off on another, promising direction. In these and other

cases, the teacher needs to exercise judgement, and may decide that the duty of acting as the devil's advocate is outweighed by another imperative at this time.

Nevertheless, the community of ethical inquiry can only claim to be achieving progress in ethical decisions if its members are enlarging their mentality through the consideration of the views of concrete others. If those others cannot be present to advance their views themselves, the teacher can do it by proxy. While the teacher cannot represent these others as well as they would themselves, students will have contact with contrary views and will both gain practice in taking them into consideration, and learn that what seems obvious to them from within their lifeworld may not seem so obvious to others, and for good reasons.

10.24 Communicative Autonomy in the Community of Inquiry

Benhabib (1992, 59n) claims that:

In education we are always counterfactually presupposing the equality and autonomy of the beings ... whose body and mind we are caring for... or training. When this counterfactual presupposition of equality, certainly not an equality of ability but one of claims, fails then we have poor pedagogics.

The presupposition of an equality of claims to which she refers seems quite reasonable. It is, if I interpret it aright, merely another way of asserting Habermas' presuppositional rationality, or the need for a procedural value of respect. However, despite the fact that she fails to expand on it, her original claim extends not just to equality but also to autonomy. In this section, I wish to address the place of autonomy in the classroom community of ethical inquiry.

In §4.23, I developed the idea of communicative autonomy, claiming that it is "the ability of the self to engage fully in the ongoing conversational narrative of humanity." This, I said, is a developmental notion, that both develops only through, and retains as an essential part, engagement with others. It is a counterfactual ideal, in that the word "fully" sets a standard that not even the well educated can meet, for the range of the conversational narrative of humanity is beyond any one individual, and McPeck's point (§10.21) that the application of critical thinking is, in detail, context-specific, means that none of us has the wherewithal to participate *fully* in all conversations.

Can teachers then, counterfactually presuppose the autonomy of the students in their classes? My answer is that they cannot. If they were to do so, then they would be led to conceiving of their task as excluding the development of communicative autonomy and the underlying capacities of reasonableness. A presupposition of autonomy would sharpen the fear of indoctrination, so that teachers would not be in a position to inculcate the necessary values of respect (§9.4). It would lead teachers to eschew the

wielding of the important strategic element to pedagogic action (§4.141). In short, it would undermine the educative endeavour.

Presupposing autonomy would be equivalent to presupposing an equality of ability, a notion that Benhabib rightly rules out. Presupposing equality of claims is a different matter. The claims of all human beings to respect are equal, whether they have it or not. A claim to respect does not develop throughout life, although a need for greater respect obviously varies with the situation in which a person is found. What this suggests is that all persons have an equality of *claim* to autonomy: a claim to a place in the conversation. This is one of the key purposes of education, and it is an important reason why one of the Universal Human Rights declared by the United Nations is the right to an education. For while it is certainly true that an education (in the institutional sense) is not the only way for an individual to develop reasonableness and autonomy (a sufficiently rich upbringing would probably suffice), it is also true that the intellectually impoverished upbringing many would receive without institutional education would prevent them from developing communicative autonomy to any great degree.

So, for some children, institutional education is not necessary for communicative autonomy, while for many others it is necessary. Even for these others, however, it is not sufficient. The nature of the education is of vital importance. I have argued throughout this thesis that the development of reasonableness is the central part of education that leads to communicative autonomy (Chapter Three), and that one of the most effective ways of developing reasonableness is through engagement in the community of inquiry. The community of inquiry is a construct of all those who participate in it, but the teacher has the most important role in this. The responsibility for looking after the interests of the community itself rests on the teacher's head. The students, as they become more reasonable and more autonomous, can begin to take on some of this responsibility, but the final responsibility, even if it reduces to merely keeping a watchful eye, always remains the teacher's.

This responsibility is exercised through the practice of pedagogical action (§4.141). The teacher (or other person placed in a position of pedagogic power) has a justified reason for engaging in the strategic elements of pedagogic actions within the community, but with this licence comes the responsibility to justify this open strategic action if and when called upon to do so, either by students within the community of inquiry or adults within the wider outside community. The justification succeeds if the teacher is able to show how those actions contribute to the development of the reasonableness and hence autonomy of the students.

My Vygotskian account of learning (§3.23) applies to the development of autonomy. Students learn through engaging jointly in the actions that we wish them to later be able to do individually. In the community of inquiry, then, teachers need to be able to scaffold the inquiry so that students participate at a higher level of autonomy than that of which they would otherwise have been capable. Here is Kant's pedagogical paradox again (§4.13) - teachers have to construct situations in which students are led to act autonomously in order to make them autonomous. But it is defused by the realization that even the endpoint - full autonomy - cannot be exercised without interaction with others, even if this is at times only the implicit interaction of internal dialogue.

10.3 Classroom Communities of Inquiry

In Part III, I have turned from the broader philosophical investigation that characterized Parts I and II, and applied those results to the classroom. We have seen that having a clearer idea of the aims of moral education, in terms of the enabling capacities of a multi-aspectual reasonableness, a vision of an interdependent autonomy, and a rich account of a meta-ethical theory to provide a basis for my characterization of an ethical person, leads to many specific conclusions as to how a classroom ought to be organized in a moral education program. In the Afterword to follow, I will summarize the main themes of this thesis.

Chapter Eleven: Afterword

I have argued that the community of ethical inquiry provides an excellent means for preparing students to be reasonable, knowledgeable and virtuous moral agents. Moral agents have many other attributes, and other strands of a moral education package are required to assist children to develop them. Many of these additional attributes are also strengthened in the community of ethical inquiry, but I do not claim that the community of ethical inquiry is sufficient to establish them on its own.

The community of ethical inquiry assists in a moral education program in three broad ways: by helping to develop reasonableness - one of the most important preconditions of thoughtful and reflective moral agency - and hence autonomy; by enhancing, through ethical inquiry, the reflective consideration of what it is that ethical living requires; and by providing the conditions in which these first two considerations can be linked to the way the students' lives are lived.

In developing the philosophical positions upon which my analysis of moral education is based, I have repeatedly drawn attention to the fact that an adequate philosophical account needs to draw on the realization that persons are always becoming-persons. We are never fully reasonable, fully autonomous, fully moral. I have been critical of philosophical systems that take rational, fully-formed adults as their starting point. An adequate philosophical account, I maintain, must take the philosophy of childhood seriously, for in considering children, we come face-to-face with the time in our lives when we are most clearly becoming-persons.

Underlying the ability to develop morally, I argued in Chapter Two, are a suite of biological presuppositions. Infants are, amongst other things, predisposed to recognize others as agents, to feel empathy with others, and to be able to develop certain kinds of reasoning. These predispositions provide both positive impulses to development, and constraints on the possible outcomes of development. They do not, however, determine the outcomes, which rely heavily on environmental (physical and social) and historical factors. It is because of this that education can be such an important factor in the formation of persons. Both these biological roots, and the child's total history provide the "raw material" on which the community of ethical inquiry can build.

11.1 Reasonableness and Autonomy

The community of inquiry provides an environment in which to develop further a self-correcting reasonableness (§3.1). Though reasonableness is built through many other

interactions in the lifeworld and even other school activities, the community of inquiry offers some features to enhance this process. Since reasonableness is acquired through the Vygotskian appropriation of jointly constructed competencies (§3.23), the teacher in the community of inquiry can deliberately construct, via pedagogical action (§4.141), the conditions for a rich interplay of reasonableness within the discussions of the community. Together with the reflexive turning of this reasonableness onto aspects of itself, the community of inquiry enables students to become more able to inquire into their developing moral understanding and commitments.

Moral agency requires more than the enhancement of “straight” as opposed to “crooked” thinking (Thouless 1946). Though important, the critical aspect of thinking (§3.121) is, I have argued, just one of five aspects to reasonableness. All the aspects play a role in thinking, especially moral thinking. The multi-aspectual account of reasonableness that I have developed (§3.12) highlights certain other preconditions for moral agency. Moral imagination plays a central role in our ethical lives (Johnson 1993), and this is well served by the creative aspect to thinking (§3.122). The committed aspect of reasonableness supports the connection of ethical thinking to interactions with, and connections to, others within a society, and plays an important role in committing persons to link their judgements to their actions (§3.123). The ability to read situations for their moral salience, and to bring to bear appropriate modes of thinking is augmented by the development of the contextual aspect of reasonableness (§3.124). Finally, students’ ability to understand the way they are physically situated within the world, interacting with it, and using embodied metaphors (such as high/low and inside/outside) which are so central to much of our moral deliberation, can be extended by attention to the embodied aspect of reasonableness (§3.125). These aspects of reasonableness, separable analytically, are intertwined in any actual thinking (§3.126)

The rich nature of the community of inquiry, including its embedding in narrative structures, its open ended character and its drawing on the life experiences of the participants, make it a more powerful vehicle for the improvement of the multiple aspects of reasonableness than many other classroom techniques. The community of inquiry is a means of engaging students in critical discourse, drawing on the competencies of lifeworld communicative action to enter into philosophical inquiry (§4.123 and §7.3). Guiding the teacher’s actions in managing the community of inquiry is the broad virtue of respect (§6.24), and the concept of an ideal speech situation (§4.124) acts as a regulative ideal to which the teacher can appeal in assessing whether respect is being instantiated within the community of inquiry.

In developing their reasonableness, students are becoming more communicatively autonomous (§3.3 and §4.2): able to act for reasons which they are able to publicly express and defend, and modify for good reasons, when called upon to do so.

However, if we take for the moment the process/content distinction seriously (and we have seen in §4.14 and §9.4 that there are reasons for being cautious in this), then the reasonableness developed through participation in a community of inquiry provides a vital process component for ethical living. For the ethical content, I will now turn my attention to the community of *ethical* inquiry.

11.2 Ethical Inquiry

Despite the conclusions of formalist theories of moral development, it is not sufficient that a moral person be just a good reasoner, able to adjudicate conflicting claims using the right type of reasoning. Nor is the development of imagination, able to be turned on moral situations, sufficient. Matthews (1994, 62-65, see §5.1) reminds us that to develop morally also involves an expanding situated knowledge of moral paradigms, the ability to define more subtly the characteristics of complex moral terms, and an increasing ability to judge whether borderline cases fall under a particular moral term. Thus, our meta-ethical target within the community of ethical inquiry must be multi-dimensional.

I have argued that virtue ethics tells us that becoming a morally good person requires the development of morally good habits, and that good moral habits cannot be unreflective (§5.2). Discourse ethics has provided me with a theoretical basis for sharpening the meaning of reflection, and arguing that it takes place in communicative action and (especially) practical discourse (Chapters Six and Seven). Further, Vygotskian engagement in practical discourse is what makes it possible for the individual to reflect ethically (§3.23).

In order to facilitate the development of ethical reflection, there is a need for more competent persons to engage in pedagogic action which, through a considered mix of open strategic action, communicative action and practical discourse, inducts younger persons into the capabilities needed for moral reflection and moral communicative autonomy (§4.141). Unavoidably underlying this pedagogic action we already find a substantive moral commitment - respect - which must be inculcated through the practice of practical discourse (§6.1).

None of this is to argue that membership of a classroom community of ethical inquiry is essential to the production of ethically good persons. Such a position would be absurd, given the recent introduction of such communities, since it would imply that very few persons are ethically good. However, in Chapter Nine, I have argued that,

because such communities formalize and concentrate the elements of communicative action and practical discourse that are most powerful in the construction of moral communicative autonomy, and because teachers are explicitly delegated pedagogic power to do so, then engagement in communities of ethical inquiry will lead to the development of persons with greater moral capabilities.

The classroom community of inquiry must therefore at regular intervals take ethical issues as its substantive focus. Within the communal inquiry, we have seen (§10.23) that both more abstract considerations of the justification of moral principles and terms, and more concrete considerations of their application in particular situations, need to be addressed, and that they should be linked together recursively. The trigger text can assist in this process by containing both specifically ethical situations and concerns, and episodes in which the characters engage in high-level conceptual puzzlement and clarification. It is then up to the teacher, acting pedagogically, to ensure that the community inquires in a similar fashion.

The mere discussion of substantive ethical matters, however, might not be seen as leading students to proper ethical knowledge and beliefs. The matter of normative correctness is highly contentious, but I have shown in §6.22 and §7.23 that discourse ethics can add power to the virtue ethics account of moral progress, even if we are not able to guarantee once-and-for-all right answers. In particular, students can gain a greater appreciation of the differences between moral issues which are more of the rightish or the goodish variety - that is, more concerned with widespread questions of justice, or personally important issues of the good life. In §10.23, I was able to apply these conclusions to the classroom community of ethical inquiry, identifying implications to help guide the teacher's pedagogical action and conduct of the discussion.

However, as I discussed in §9.21, even a well developed reasonableness and a sophisticated ethical knowledge is not sufficient to guarantee that students will live an ethical life, if their judgements and actions are divorced from one another. The final claim of this thesis is that the community of ethical inquiry does have resources that assist in marrying the two.

11.3 Lived Lives

To turn to the lived lives of students is to remind ourselves of the broader picture. A single pedagogical method used in a single classroom, no matter how good it is, will not be enough to guarantee the formation of moral persons. Even the efforts of the school as a whole may well not be enough. The other influences on becoming-persons are also very powerful. Yet this does not absolve schools and teachers of a

responsibility to do the best they can. In this thesis, I have argued that the best, or at least a very good, method available to teachers is the community of ethical inquiry.

The community of ethical inquiry, however, takes place in a restricted context - the classroom. It involves not just talk, but also action, because speech is itself a type of action, and because the participants in the community are involved in interpersonal interaction. These actions, however, like the community, are also restricted in their scope. Becoming-persons, and the persons into which they will grow, will need to act in much broader contexts. Given the difficulties I have identified (§3.1245 and §10.221) in transfer across contexts, we have some grounds for concern that even the most virtuous member of the community may not transfer that virtue into the remainder of their life.

I have argued that, although we do not have a guarantee, we can have grounds for optimism that the community of ethical inquiry will encourage transfer. The community of ethical inquiry offers a rich narrative environment that links philosophical considerations to the lives of persons in two ways. Firstly, the texts provide models of persons who integrate their deliberations into their lives (§10.1). Secondly, the judicious encouragement of anecdotal reflection within the inquiry, when linked to the more abstract contemplations of the community, provides a path for the incorporation of the fruit of the community's meditations into the students' personal lives (§10.223).

This is not the only contribution the community of ethical inquiry can make, for the points in the previous paragraph concentrate largely on the inquiry rather than the community. This latter aspect connects each member of the community to the other members. The community is, within a restricted context, a microcosm of larger communities. Each member learns within this community what it is to act with respect. Though there are, as I have noted in §10.221, problems with transfer from one context to another, the reflexive nature of the community of ethical inquiry does encourage such transfer.

Nevertheless, we cannot escape the conclusion that I have been urging throughout. The community of ethical inquiry has special virtues, and provides particular advantages, in a program of moral education. But it cannot form the totality of that program. Moral education requires a broader, multi-dimensional approach, within which the community of ethical inquiry holds an important, but not a sole, place.

11.4 Conclusion

In Chapter One, I referred to the state of global society and the need for each generation to bring up the next in such a way as to maintain what is morally worthwhile in society, to alter what is not, and to be able to both plan for, and react to, the rapid changes of society in morally acceptable ways. This is a complex task, and it requires multiple inputs from many social actors and sectors of society.

Central to these endeavours are two particular institutions: the family and the education system. Within the education system, I have again narrowed my focus, to the classroom itself, and to the place of a particular pedagogical methodology: the community of ethical inquiry. I have made a series of recommendations to teachers as to how such a community should be set up, and what the benefits will be if this is done.

My focus has been on the teacher and the classroom, but I believe that parents too could benefit from my work. For they too engage in pedagogical action with their children. Their actions are vital in the establishment of good moral habits which, as we have seen, also requires the habituation of reasonableness. If moral education is to be successful, and produce citizens who are reasonable, autonomous and ethical, then families and schools will need to assist each other in equipping children to inquire communally into ethical matters.

Bibliography

- AEC (Australian Education Council) (1994). *A Statement on Studies of Society and Environment for Australian Schools*. Melbourne, Curriculum Corporation.
- Alexandra, P. A. and J. E. Judy (1988). The Interaction of Domain-Specific and Strategic Knowledge in Academic Performance. *Review of Educational Research* 58 (4), 375-404.
- Allison, H. E. (1996). *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy*. Cambridge, Cambridge University Press.
- Apel, K.-O. (1980). The A Priori of the Communication Community and the Foundations of Ethics. In K.-O. Apel (Ed.), *Towards a transformation of philosophy* London: Routledge & Kegan Paul.
- Aristotle (1980). *The Nicomachean Ethics*. Oxford, Oxford University Press.
- Arnold, M. (1888). *Essays on Criticism (Second series)*. London, Macmillan.
- Baier, A. (1985). *Postures of the Mind: Essays on Mind and Morals*. London, Methuen.
- Baier, A. (1994). *Moral Prejudices: Essays on Ethics*. Cambridge, MA, Harvard University Press.
- Baier, K. (1971). Ethical Pluralism and Moral Education. In C. M. Beck, B. S. Crittenden and E. V. Sullivan (Ed.), *Moral Education: Interdisciplinary Approaches* (93-112). New York: Newman Press.
- Barnes, D. and F. Todd (1977). *Communication and Learning in Small Groups*. London, Routledge.
- Baron, M. (1985). The Ethics of Duty/Ethics of Virtue Debate and its Relevance to Educational Theory. *Educational Theory* 35 (2), 135-149.
- Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA, MIT Press.
- Beardsley, M. C. and E. L. Beardsley (1965). *Philosophical Thinking: An Introduction*. New York, Harcourt, Brace and World.
- Beck, C. M., B. S. Crittenden, et al. (Ed.) (1971). *Moral Education: Interdisciplinary Approaches*. New York: Newman Press.
- Benhabib, S. (1992). *Situating the Self: Gender, Community and Postmodernism in Contemporary Ethics*. Cambridge, Polity Press.
- Biesta, G. (1995). Pragmatism as a Pedagogy of Communicative Action. *Studies in Philosophy and Education* 13 , 273-290.
- Billig, M. (1987). *Arguing and Thinking: A Rhetorical Approach to Social Psychology*. Cambridge, Cambridge University Press.

- Blake, N. (1995). Ideal Speech Conditions, Modern Discourse and Education. *Journal of Philosophy of Education* 29 (3), 355–367.
- Boostrom, R. (1992). *Developing Creative & Critical Thinking : An Integrated Approach*. Lincolnwood, Ill, National Textbook Co.
- Braaten, J. (1991). *Habermas's Critical Theory of Society*. Albany, NY, State University of New York Press.
- Braaten, J. (1995). From Communicative Rationality to Communicative Thinking: A Basis for Feminist Theory and Practice. In J. Meehan (Ed.), *Feminists Read Habermas: Gendering the Subject of Discourse* (139–162). New York: Routledge.
- Brabeck, M. M. (1983). Critical Thinking Skills and Reflective Judgement Development: Redefining the Aims of Higher Education. *Journal of Applied Developmental Psychology* 4 (1), 23–24.
- Brand, A. (1990). *The Force of Reason: An Introduction to Habermas' Theory of Communicative Action*. North Sydney, Allen & Unwin.
- Broad, C. D. (1978). *Kant: An Introduction*. Cambridge, Cambridge University Press.
- Brown, D. E. (1991). *Human Universals*. Philadelphia, Temple University Press.
- Bruner, J. (1974). *Beyond the Information Given: Studies in the Psychology of Knowing*. London: Allen & Unwin.
- Bruner, J. (1983). *Child's Talk: Learning to Use Language*. Oxford, Oxford University Press.
- Bruner, J. S. (1986). *Actual Minds, Possible Worlds*. Cambridge, MA, Harvard University Press.
- Burnyeat, M. F. (1980). Aristotle on Learning to be Good. In A. O. Rorty (Ed.), *Essays on Aristotle's Ethics* (69–92). Berkeley: University of California Press.
- Cahoone, L. (Ed.) (1996). *From Modernism to Postmodernism*. Oxford: Blackwell.
- Cam, P. (1994). A Philosophical Approach to Moral Education. *Critical & Creative Thinking* 2 (2), 19–26.
- Cam, P. (1995). Against Indoctrination: Response to Davson-Galle. *Critical & Creative Thinking* 3 (1), 38–40.
- Cam, P. (1997). *Thinking Stories* 3. Sydney, Hale & Iremonger.
- Caygill, H. (1995). *A Kant Dictionary*. Oxford, Blackwell.
- Chan, H.-M. (1998). *Whose Justice? A Metaphilosophical Reflection on the Study of Ethics*. Paper delivered at the Australasian Society of Asian and Comparative Philosophy Conference, University of New South Wales (July 1998).
- Cherniak, C. (1986). *Minimal Rationality*. Cambridge, MA, MIT Press.
- Chomsky, N. (1968). *Language and Mind*. New York, Harcourt Brace Jovanovich.

- Clark, A. (1998). *Being There: Putting Brain, Body and World Together Again*. Cambridge, MA, MIT Press.
- Claxton, G. (1991). *Educating the Inquiring Mind*. London, Harvester Wheatsheaf.
- Code, L. (1991). *What Can She Know?: Feminist Theory and the Construction of Knowledge*. Ithaca, Cornell University Press.
- Cohen, J. L. (1995). Critical Social Theory and Feminist Critiques: The Debate with Jürgen Habermas. In J. Meehan (Ed.), *Feminists Read Habermas: Gendering the Subject of Discourse* (57–90). New York: Routledge.
- Cosmides, L. and J. Tooby (1994). Beyond Intuition and Instinct Blindness: Towards an Evolutionarily Rigorous Cognitive Science. *Cognition* 50, 41-77.
- Cresswell, R. and P. Hobson (1994). Contested Values and Philosophy for Children in a Pluralist Democratic State. *Critical & Creative Thinking* 2 (2), 29-35.
- Damasio, A. R. (1996). *Descartes' Error*. London, Papermac.
- Damon, W. (1988). *The Moral Child*. New York, Free Press.
- Damon, W. (1996). The Lifelong Transformation of Moral Goals Through Social Influence. In P. B. Baltes and U. M. Staudinger (Ed.), *Interactive Minds: Life-Span Perspectives on the Social Foundations of Cognition* (198–220). New York: Cambridge University Press.
- de Bono, E. (1985). The CoRT Thinking Program. In J. Segal, S. Chipman and R. Glaser (Ed.), *Thinking and Learning Skills* Hillsdale NJ: Laurance Erlbaum.
- de Sousa, R. (1987). *The Rationality of Emotion*. Cambridge MA, MIT Press.
- Descartes, R. (1984). Meditations on First Philosophy. In J. Cottingham, R. Stoothoff and D. Murdoch (Ed.), *The Philosophical Writings of Descartes* (1-62). Cambridge: Cambridge University Press.
- Dewey, J. (1966). *Democracy and Education*. New York, The Free Press.
- di Giovanni, G. (1992). The First Twenty Years of Critique: The Spinoza Connection. In P. Guyer (Ed.), *The Cambridge Companion to Kant* (417-448). Cambridge: Cambridge University Press.
- Dillon, J. T. (1994). *Using Discussion in Classrooms*. Buckingham and Philadelphia, Open University Press.
- diSessa, A. A. (1988). Knowledge in Pieces. In G. Forman and P. B. Pufall (Ed.), *Constructivism in the Computer Age* (49-70). Hillsdale, NJ: Lawrence Erlbaum.
- Donaldson, M. (1978). *Children's Minds*. London, Fontana.
- Dreyfus, H. L. (1992). *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge, MA, MIT Press.
- Edwards, D. and N. Mercer (1987). *Common Knowledge: The Development of Understanding in the Classroom*. London, Methuen.

- Elias, J. L. (1989). *Moral Education: Secular and Religious*. Malabar, FL, Robert E. Kreiger.
- Ennis, R. H. (1987). A Taxonomy of Critical Thinking Dispositions and Abilities. In J. B. Baron and R. J. Sternberg (Ed.), *Teaching Thinking Skills: Theory and Practice* (9-26). New York: Freeman.
- Facione, P. (Ed.) (1989). *Report on Critical Thinking*. Newark, DE: American Philosophical Association Subcommittee on Pre-College Philosophy, University of Delaware.
- Feyerabend, P. (1975). *Against Method: Outline of an Anarchistic Theory of Knowledge*. London, New Left Books.
- Flanagan, O. (1991). *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge, MA, Harvard University Press.
- Forman, E. A. and C. B. Cazden (1985). Exploring Vygotskian Perspectives in Education: The Cognitive Value of Peer Interaction. In J. V. Wertsch (Ed.), *Culture, Communication and Cognition* (323-347). Cambridge: Cambridge University Press.
- Fosnot, C. T. (1993). Rethinking Science Education: A Defense of Piagetian Constructivism. *Journal of Research in Science Teaching* 30 (9), 1189-1201.
- Foucault, M. (1991). Truth and Power. In P. Rabinow (Ed.), *The Foucault Reader* (51-75). London: Penguin.
- Fowler, H. W. and E. Gowers (1965). *A Dictionary of Modern English Usage*. Oxford, Oxford University Press.
- Fraenkel, J. (1977). *How to Teach about Values: An Analytic Approach*. Englewood Cliffs, NJ, Prentice-Hall.
- Frohlich, N. and J. A. Oppenheimer (1992). *Choosing Justice: An Experimental Approach to Ethical Theory*. Berkeley, CA, University of California Press.
- Gadamer, H.-G. (1975). *Truth and Method*. New York, Seabury Press.
- Gardner, H. (1985). *Frames of Mind: The Theory of Multiple Intelligences*. New York, Basic Books.
- Gardner, S. (1995). Inquiry is No Mere Conversation. *Critical & Creative Thinking* 3 (2), 38-49.
- Garfield, J. L. (1997). Choosing Coercion and Subordination - A Preliminary Moral Study of Academic Life. In S. Hellsten, M. Kopperi and O. Loukola (Ed.), *Taking the Liberal Challenge Seriously* (305-317). Aldershot: Ashgate.
- Gauthier, D. P. (1971). Moral Action and Moral Education. In C. M. Beck, B. S. Crittenden and E. V. Sullivan (Ed.), *Moral Education: Interdisciplinary Approaches* (138-146). New York: Newman Press.
- Gazzard, A. (1988). Evidence of Effectiveness of the Philosophy for Children Program - Quantitative Studies - 1987-1988 in *Philosophy for Children: Where Are We Now.... Thinking* 8 (Supplement two), S13-S14.

- Gilligan, C. (1982). *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA, Harvard University Press.
- Glaser, J. (1998). *Reasoning, Critical Thinking and the Critical Person: Towards a Dialogical Theory of Critical Thinking*. Unpublished PhD Thesis. Philosophy Department. Melbourne, University of Melbourne.
- Glaser, J. (1998). Thinking Together: Arendt's Visiting Imagination and Nussbaum's Judicial Spectatorship as Models for a Community of Inquiry. *Thinking* 14 (1), 17-23.
- Habermas, J. (1971). Vorbereitende Bemerkungen zu einer Theorie der kommunikativen Kompetenz. In J. Habermas and N. Luhmann (Ed.), *Theorie der Gesellschaft oder Sozialtechnologie* Frankfurt am Main: Suhrkamp.
- Habermas, J. (1981). *The Theory of Communicative Action*. Boston, Beacon Press.
- Habermas, J. (1990). *Moral Consciousness and Communicative Action*. Cambridge, MA, MIT Press.
- Habermas, J. (1992). *Postmetaphysical Thinking*. Cambridge, Polity Press.
- Habermas, J. (1993). *Justification and Application: Remarks on Discourse Ethics*. Cambridge, MA, MIT Press.
- Habermas, J. (1996). *Between Facts and Norms*. Cambridge, Polity Press.
- Hammer, D. (1996). Misconceptions or P-prims: How May Alternative Perspectives of Cognitive Structure Influence Instructional Perceptions and Intentions? *Journal of the Learning Sciences* 5 (2), 97-127.
- Hamston, J. (1995). *Validating Students' Voices : Critical Literacy in the Primary Classroom*. Canberra, National Languages and Literacy Institute of Australia.
- Haste, H. (1996). Communitarianism and the Social Construction of Morality. *Journal of Moral Education* 25 (1), 47-55.
- Haugeland, J. (1979). Understanding Natural Language. *Journal of Philosophy* 76 (November 1979), 619-632.
- Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*. Cambridge, MA, MIT Press.
- Havel, P. (1995). Values Framework. Perth, WA, Values Review NPDP Project.
- Heidegger, M. (1962). *Being and Time*. New York, Harper & Row.
- Hennessy, S. (1993). Situated Cognition and Cognitive Apprenticeship: Implications for Classroom Learning. *Studies in Science Education* 22 , 1-41.
- Henry, M., J. Knight, et al. (1988). *Understanding Schooling: An Introductory Sociology of Australian Education*. London, Routledge.
- Hobbes, T. (1966). Philosophical Rudiments Concerning Government and Society. In W. Molesworth (Ed.), *The English Works of Thomas Hobbes* Darmstadt: Wissenschaftliche Buchgesellschaft.

- Horster, D. (1992). *Habermas: An Introduction*. Philadelphia, Pennbridge Books.
- Howe, A. C. (1996). Development of Science Concepts in a Vygotskian Framework. *Science Education* 80 (1), 35-51.
- Hoy, D. C. (1986). Power, Repression, Progress: Foucault, Lukes, and the Frankfurt School. In D. C. Hoy (Ed.), *Foucault: a critical reader* (123-143). Oxford: Blackwell.
- Hull, R. (1985). *The Language Gap: How Classroom Dialogue Fails*. London, Methuen.
- Hume, D. (1739/1962). *A Treatise of Human Nature: Book One*. London, Collins/Fontana.
- Hume, D. (1740/1972). *A Treatise of Human Nature: Books Two and Three*. London, Fontana.
- Hume, D. (1955). An Enquiry Concerning the Principles of Morals. In C. W. Hendel (Ed.), *Hume Selections* (194-252). New York: Charles Scribner's Sons.
- Hutchinson, D. S. (1995). Ethics. In J. Barnes (Ed.), *The Cambridge Companion to Aristotle* (195-232). Cambridge: Cambridge University Press.
- Inhelder, B., H. Sinclair, et al. (1974). *Learning and the Development of Cognition*. London, Routledge & Kegan Paul.
- Jackson, M. (1994). Approaches to Learning Ethics. In N. Preston (Ed.), *Ethics for the Public Sector: Education and Training* (179-190). Sydney: Federation Press.
- Jaggar, A. M. (1992). Love and Knowledge: Emotion in Feminist Epistemology. In E. D. Harvey and K. Okruhlik (Ed.), *Women and reason* (115-142). Ann Arbor: University of Michigan Press.
- Jepson, R. W. (1952). *Clear Thinking*. London, Longmans, Green.
- Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago, University of Chicago Press.
- Johnson, M. (1993). *Moral Imagination: Implications of Cognitive Science for Ethics*. Chicago, University of Chicago Press.
- Kant, I. (1965). *Critique of Pure Reason*. New York, St Martin's Press.
- Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA, MIT Press.
- Kivelä, A. (1998). *Pedagogical Theory Gets Nowhere: Notes on H. Peukert's Article "Basic Problems of a Critical Theory of Education"*. University of Oulu. URL: <http://wwwedu.oulu.fi/homepage/yptr/peuk.htm> (accessed March 1998).
- Kivelä, A., J. Peltonen, et al. (1998). *Towards a Theory of Pedagogic Action*. University of Oulu. URL: <http://wwwedu.oulu.fi/homepage/yptr/toward.htm> (accessed March 1998).

- Kohlberg, L. (1971). Stages of Moral Development as a Basis for Moral Education. In C. M. Beck, B. S. Crittenden and E. V. Sullivan (Ed.), *Moral Education: Interdisciplinary Approaches* (23-92). New York: Newman Press.
- Kohlberg, L. (1987). *Child Psychology and Childhood Education*. New York, Longman.
- Kohlberg, L., C. Levine, et al. (1983). *Moral Stages: A Current Formulation and Response to Critics*. Basel, Karger.
- Kohlberg, L. and J. Wertsch (1987). Language and the Development of Thought. In L. Kohlberg (Ed.), *Child Psychology and Childhood Education* New York: Longman.
- Korsgaard, C. M. (1996). *Creating the Kingdom of Ends*. Cambridge, Cambridge University Press.
- Lankshear, C. and P. McLaren (Ed.) (1993). *Critical Literacy : Politics, Praxis, and the Postmodern*. Albany, NY: State University of New York Press.
- Larkin, J., J. McDermott, et al. (1980). Expert and Novice Performance in Solving Physics Problems. *Science* 208 , 1335-1342.
- Lawson, A. E. (1990). Use of Reasoning to a Contradiction in Grade 3 to College. *Journal of Research in Science Teaching* 27 (6), 541-551.
- Lemke, J. L. (1990). *Talking Science: Language, Learning and Values*. Norwood, NJ, Ablex.
- Lewis, C. S. (1946). *The Abolition of Man*. London, Geoffrey Bles.
- Lipman, M. (1983). *Lisa*. Upper Montclair, NJ, IAPC.
- Lipman, M. (1991). *Thinking in Education*. Cambridge, Cambridge University Press.
- Lipman, M. (1995). Caring as Thinking. *Inquiry: Critical Thinking Across the Disciplines* 15 (1), 1-13.
- Lipman, M. (1997). Why Caring Matters. Unpublished manuscript.
- Lipman, M. and A. Gazzard (1986). Evidence of Effectiveness - Quantitative Studies in Philosophy for Children: Where Are We Now.... *Thinking* 6 (Supplement), S6-S8.
- Lipman, M. and A. M. Sharp (1985). *Ethical Inquiry: Instruction Manual to Accompany Lisa*. Lanham, MD, University Press of America.
- Lipman, M., A. M. Sharp, et al. (1980). *Philosophy in the Classroom*. Philadelphia, Temple University Press.
- Lipman, M., A. M. Sharp, et al. (1984). *Philosophical Inquiry: An Instructional Manual to Accompany Harry Stottlemeier's Discovery*. Lanham, MD, University Press of America.
- Lipman, M. and L. Splitter (1992). *Harry Stottlemeier's Discovery*. Melbourne, ACER.

- Lloyd, G. (1984). *The Man of Reason: "Male" and "Female" in Western Philosophy*. London, Methuen.
- Loubser, J. J. (1971). The Contribution of Schools to Moral Development: A Working Paper in the Theory of Action. In C. M. Beck, B. S. Crittenden and E. V. Sullivan (Ed.), *Moral Education: Interdisciplinary Approaches* (147-179). New York: Newman Press.
- Lutz, C. (1988). *Unnatural Emotions*. Chicago, Chicago University Press.
- Lynch, P. P. and B. L. Jones (1995). Students' Alternative Frameworks: Towards a Linguistic and Cultural Interpretation. *International Journal of Science Education* 17 (1), 107-118.
- Lyotard, J. F. (1984). *The Postmodern Condition: A Report on Knowledge*. Manchester, Manchester University Press.
- MacColl, S. (1992). Conversation and Dialogue - Not Just What You Say but How You Say It. In R. Reed (Ed.), *When We Talk: Essays on Classroom Conversation* (75-85). Fort Worth: Analytic Teaching Press.
- Macintyre, S., K. Boston, et al. (1994). *Whereas the People...: Civics and Citizenship Education*. Canberra, Australian Government Publishing Service.
- Masschelein, J. (1991). *Kommunikatives Handeln und pädagogisches Handeln. Die Bedeutung der Habermaschen kommunikationstheoretischen Wende für die Pädagogik. [Communicative Action and Pedagogic Action: The Significance of Habermasian Communications Theoretical Change for Pedagogy]*. Weinheim, Deutscher Studien Verlag.
- Masschelein, J. (1991). The Relevance of Habermas' Communicative Turn. *Studies in Philosophy and Education* 11, 95-111.
- Matthews, G. (1980). *Philosophy and the Young Child*. Cambridge, Harvard University Press.
- Matthews, G. (1994). *The Philosophy of Childhood*. Cambridge, MA, Harvard University Press.
- McMillan, J. H. (1987). Enhancing College Students' Critical Thinking: A Review of Studies. *Research in Higher Education* 26 (1), 3-19.
- McPeck, J. E. (1981). *Critical Thinking and Education*. Oxford, Martin Robertson.
- McPeck, J. E. (1990). *Teaching Critical Thinking*. London, Routledge.
- Mead, G. H. (1982). *The Individual and the Social Self*. Chicago, Chicago University Press.
- Meadows, S. (1988). Piaget's Contribution to Understanding Cognitive Development: An Assessment for the late 1980s. In K. R. S. Sheldon (Ed.), *Cognitive Development to Adolescence* (19-31). Hove, Sussex: Open University/Lawrence Erlbaum Associates.
- Meehan, J. (1995). Autonomy, Recognition and Respect: Habermas, Benjamin, and Honneth. In J. Meehan (Ed.), *Feminists Read Habermas: Gendering the Subject of Discourse* (231-246). New York: Routledge.

- Meehan, J. (Ed.) (1995). *Feminists Read Habermas: Gendering the Subject of Discourse*. New York: Routledge.
- Melden, A. I. (1971). Moral Education and Moral Action. In C. M. Beck, B. S. Crittenden and E. V. Sullivan (Ed.), *Moral Education: Interdisciplinary Approaches* (115-137). New York: Newman Press.
- Mill, J. S. (1965). Utilitarianism. In J. B. Schneewind (Ed.), *Mill's Ethical Writings*. New York: Collier.
- Millar, R. and W. Kragh (1994). Alternative Frameworks or Context-Specific Reasoning? Children's Ideas about the Motion of Projectiles. *School Science Review* 75 (272), 27-34.
- Morgan, J. (1996). A Defense of Autonomy as an Educational Ideal. *Journal of Philosophy of Education* 30 (2), 239-252.
- Morgan, W. (Ed.) (1996). *Critical Literacy: Readings and Resources [kit]*. Norwood, S. Aust.: Australian Association for the Teaching of English.
- Noddings, N. (1996). On Community. *Educational Theory* 46 (3), 245-267.
- Norris, S. (1990). Thinking about Critical Thinking: Philosophers Can't Go It Alone. In J. E. McPeck (Ed.), *Teaching Critical Thinking* (67-74). London: Routledge.
- Nussbaum, M. (1995). *Poetic Justice*. Boston, Beacon Press.
- Nussbaum, M. (1997). Emotions as Judgements of Value and Importance. In P. Bilimoria and J.N.Mohanty (Ed.), *Relativism, Suffering and Beyond: Essays In Memory of Bimal K. Matilal* (231-251). Oxford: Oxford University Press.
- Nussbaum, M. C. (1994). *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*. Princeton, Princeton University Press.
- O'Neill, O. (1989). *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge, Cambridge University Press.
- Osborne, J. F. (1996). Beyond Constructivism. *Science Education* 80 (1), 53-82.
- Pascarella, E. T. and P. T. Terenzini (1991). *How College Affects Students: Findings and Insights from Twenty Years of Research*. San Francisco, Jossey-Bass.
- Passmore, J. (1972). On Teaching to be Critical. In R. F. Dearden, P. H. Hirst and R. S. Peters (Ed.), *Education and Reason: Part 3 of "Education and the Development of Reason"* (25-43). London: Routledge & Kegan Paul.
- Paul, R. (1990). McPeck's Mistakes. In J. E. McPeck (Ed.), *Teaching Critical Thinking* (102-111). London: Routledge.
- Peirce, C. S. (1955). *Philosophical Writings of Peirce*. New York, Dover.
- Perkins, D. (1994). The Hidden Order of Open-Ended Thinking. In J. Edwards (Ed.), *Thinking: International Interdisciplinary Perspectives* (85-96). Melbourne: Hawker Brownlow.
- Perkins, D. N. and G. Salomon (1989). Are Cognitive Skills Context-Bound? *Educational Researcher* 18 (1), 16-25.

- Peters, R. S. (1966). *Ethics and Education*. London, George Allen & Unwin.
- Pettit, P. (1993). *The Common Mind*. Oxford, Oxford University Press.
- Peukert, H. (1993). Basic Problems of a Critical Theory of Education. *Journal of Philosophy of Education* 27 (2), 159-170.
- Phillips, R. (1994). A Sincere Word for the Devil's Advocate. *Critical & Creative Thinking* 2 (1), 15-20.
- Piaget, J. (1977). *The Moral Judgement of the Child*. Harmondsworth, Penguin.
- Pitt, J., C. Tugwell, et al. (1995). Not Just After Lunch on Wednesdays [kit]. Adelaide, S. Aust., Dept. for Education and Children's Services. Education Video Unit.
- Polanyi, M. (1958). *Personal Knowledge*. Henley, Routledge & Keegan Paul.
- Pontecorvo, C. and H. Girardet (1993). Arguing and Reasoning in Understanding Historical Topics. *Cognition and Instruction* 11 (3 & 4), 365-395.
- Popper, K. (1962). *The Open Society and its Enemies*. London, Routledge & Kegan Paul.
- Postman, N. and C. Weingartner (1969). *Teaching as a Subversive Activity*. New York, Delacorte Press.
- Power, F. C., A. Higgins, et al. (1989). *Laurance Kohlberg's Approach to Moral Education*. New York, Columbia University Press.
- Pritchard, M. S. (1992). Moral Education: From Aristotle to Harry Stottlemeier. In A. M. Sharp and R. F. Reed (Ed.), *Studies in Philosophy for Children: Harry Stottlemeier's Discovery* (15-31). Philadelphia: Temple University Press.
- Pritchard, M. S. (1996). *Reasonable Children: Moral Education and Moral Learning*. Lawrence, KS, University Press of Kansas.
- Project, S. C. N. H. (1970). *The Humanities Project: An Introduction*. London, Heinemann.
- Pylyshyn, Z. W. (Ed.) (1987). *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. Norwood, NJ: Ablex.
- Rachels, J. (1993). *The Elements of Moral Philosophy*. New York, McGraw-Hill.
- Raths, L. E., M. Harmin, et al. (1978). *Values and Teaching*. Columbus, Charles E Merrill.
- Rawls, J. (1973). *A Theory of Justice*. Oxford, Oxford University Press.
- Raz, J. (1986). *The Morality of Freedom*. Oxford, Clarendon Press.
- Read, D., A. and S. B. Simon (Ed.) (1975). *Humanistic Education Sourcebook*. Englewood Cliffs, NJ: Prentice-Hall.
- Reed, R. (1992). Inventing a Classroom Conversation. In A. M. Sharp and R. Reed (Ed.), *Studies in Philosophy for Children: Harry Stottlemeier's Discovery* (158-164). Philadelphia: Temple University Press.

- Reed, R. (1992). On the Art and Craft of Dialogue. In A. M. Sharp and R. Reed (Ed.), *Studies in Philosophy for Children: Harry Stottlemeier's Discovery* (147-157). Philadelphia: Temple University Press.
- Reed, R. (1992). On the Use of Criteria: From Anecdote to Meta-Criteria. In R. Reed (Ed.), *When We Talk: Essays on Classroom Conversation* (87-96). Fort Worth: Analytic Teaching Press.
- Reichenbach, H. (1968). *The Rise of Scientific Philosophy*. Los Angeles, University of California Press.
- Reid, T. (1872). Essays on the Active Powers of the Human Mind. In S. W. Hamilton (Ed.), *The Works of Thomas Reid, D. D.* Edinburgh: Maclachlan and Stewart.
- Rogoff, B. (1984). Introduction: Thinking and Learning in Social Context. In B. Rogoff and J. Lave (Ed.), *Everyday Cognition: Its Development in Social Contexts* (1-8). Cambridge, MA: Harvard University Press.
- Rogoff, B. and W. Gardner (1984). Adult Guidance of Cognitive Development. In B. Rogoff and J. Lave (Ed.), *Everyday Cognition: Its Development in Social Contexts* (95-116). Cambridge, MA: Harvard University Press.
- Rosch, E. and B. B. Lloyd (Ed.) (1978). *Cognition and Categorization*. Hillsdale, NJ: Lawrence Erlbaum.
- Rowe, C. (1993). Ethics in Ancient Greece. In P. Singer (Ed.), *A Companion to Ethics* (121-132). Oxford: Blackwell.
- Schneewind, J. B. (1992). Autonomy, Obligation, and Virtue: An Overview of Kant's Moral Philosophy. In P. Guyer (Ed.), *The Cambridge Companion to Kant* (309-341). Cambridge: Cambridge University Press.
- Sellars, W. (1963). *Science, Perception and Reality*. London, Routledge & Kegan Paul.
- Sherman, N. (1989). *The Fabric of Character: Aristotle's Theory of Virtue*. Oxford, Clarendon Press.
- Siegel, H. (1988). *Educating Reason: Rationality, Critical Thinking and Education*. London, Routledge.
- Siegel, H. (1990). McPeck, Informal Logic, and the Nature of Critical Thinking. In J. E. McPeck (Ed.), *Teaching Critical Thinking* (75-85). London: Routledge.
- Sinclair, J. M. and R. M. Coulthard (1975). *Towards an Analysis of Discourse: The English Used by Pupils and Teachers*. Oxford, Oxford University Press.
- Singer, P. (Ed.) (1993). *A Companion to Ethics*. Oxford: Blackwell.
- Skinner, B. F. (1971). *Beyond Freedom and Dignity*. New York, Alfred A. Knopf.
- Skurzynski, G. (1980). *Honest Andrew*. New York, Harcourt Brace Jovanovich.
- Solomon, J. (1992). *Getting to Know About Energy - In School and Society*. London, Falmer.

- Solomon, J. (1994). The Rise and Fall of Constructivism. *Studies in Science Education* 23, 1-19.
- Solomon, R. (1988). On Emotions as Judgments. *American Philosophical Quarterly* 25 (2), 183-191.
- Solomon, R. C. (1983). *The Passions: The Myth and Nature of Human Emotion*. Notre Dame, University of Notre Dame Press.
- Sorabji, R. (1980). Aristotle on the Role of Intellect in Virtue. In A. O. Rorty (Ed.), *Essays on Aristotle's ethics* (201-219). Berkeley: University of California Press.
- Spelman, E. V. (1988). *Inessential Woman: Problems of Exclusion in Feminist Thought*. Boston, Beacon Press.
- Splitter, L. (1997). Some Reflections on Inquiry, Community and Philosophy. *Inquiry: Critical Thinking across the Disciplines* 17 (1), 29-39.
- Splitter, L. and A. M. Sharp (1995). *Teaching for Better Thinking: The Classroom Community of Inquiry*. Melbourne, ACER.
- Sprod, T. (1993). *Books into Ideas*. Melbourne, Hawker Brownlow.
- Sprod, T. (1994a). An Attempt to Evaluate Students' Reactions to the Implementation of P4C at Hutchins School, Tasmania. *Critical & Creative Thinking* 2 (1), 59-65.
- Sprod, T. (1994b). Cognitive Development, Philosophy and Children's Literature. *Early Child Development and Care* 107, 23-33.
- Sprod, T. (1994c). *Developing Higher Order Thinking through Whole Class Discussion in a Science Classroom*. Unpublished MSc Thesis, Department of Educational Studies, University of Oxford.
- Sprod, T. (1997a). What is a Community of Inquiry?: Consideration on an Email Discussion List. *Inquiry: Critical Thinking Across the Disciplines* 17 (1), 4-28.
- Sprod, T. (1997b). The Cape Grim and Emu Bay Incidents: An Historical Community of Inquiry. *Critical & Creative Thinking* 5 (1), 42-49.
- Sprod, T. (1997c). Improving Scientific Reasoning through Philosophy for Children: An Empirical Study. *Thinking* 13 (2), 75-80.
- Sprod, T. (1997d). "Nobody Really Knows": The Structure and Analysis of Social Constructivist Whole Class Discussions. *International Journal of Science Education* 19 (8), 911-924.
- Sprod, T. (1998). "I Can Change Your Opinion on That": Social Constructivist Whole Class Discussions and Their Effect on Scientific Reasoning. *Research in Science Education* 28 (4), 463-480.
- Stebbing, L. S. (1939). *Thinking to Some Purpose*. Harmondsworth, Pelican.
- Straughan, R. (1988). *Can We Teach Children to be Good?* Milton Keynes, Open University Press.

- Strawson, P. F. (1964). *Individuals*. London, Methuen.
- Strawson, P. F. (1974). Freedom and Resentment. In P. F. Strawson (Ed.), *Freedom and Resentment and Other Essays* (1-25). London: Methuen.
- Taylor, C. (1989). *Sources of the Self: The Making of the Modern Identity*. Cambridge, MA, Harvard University Press.
- Taylor, C. (1991). Language and Society. In A. Honneth and H. Joas (Ed.), *Communicative Action: Essays on Jürgen Habermas's The Theory of Communicative Action* (23-35). Cambridge: Polity Press.
- Taylor, C. (1995). The Dialogical Self. In R. F. Goodman and W. R. Fisher (Ed.), *Rethinking Knowledge: Reflections across the Disciplines* Albany: State University of New York Press.
- Thompson, J. B. (1981). *Critical Hermeneutics: A Study in the Thought of Paul Ricoeur and Jürgen Habermas*. Cambridge, Cambridge University Press.
- Thouless, R. H. (1946). *Straight and Crooked Thinking*. London, The English Universities Press.
- Tizard, B. and M. Hughes (1984). *Young Children Learning: Talking and Thinking at Home and School*. London, Fontana.
- Tytler, R. (1994). Consistency of Children's Use of Scientific Conceptions: Problems with the Notion of "Conceptual Change". *Research in Science Education* 24 , 338-347.
- van Gelder, T. (1998). *Critical Thinking: Some Myths and Misconceptions*. School of Philosophy Research Seminar, University of Tasmania (27 May 1998).
- van Manen, M. (1991). *The Tact of Teaching: The Meaning of Pedagogical Thoughtfulness*. Albany, SUNY Press.
- Vygotsky, L. S. (1962). *Thought and Language*. Cambridge, MIT Press.
- Vygotsky, L. S. (1981). The Genesis of the Higher Mental Functions. In J. V. Wertsch (Ed.), *The Concept of Activity in Soviet Psychology* Armonk, NY: M. E. Sharpe.
- Warnke, G. (1995). Discourse Ethics and Feminist Dilemmas of Difference. In J. Meehan (Ed.), *Feminists Read Habermas: Gendering the Subject of Discourse* (247-262). New York: Routledge.
- Wartenberg, T. E. (1990). *The Forms of Power: From Domination to Transformation*. Philadelphia, Temple University Press.
- Wells, G., G. L. M. Chang, et al. (1990). Creating Classroom Communities of Literate Thinkers. In S. Schlomo (Ed.), *Cooperative Learning: Theory and Research* New York: Praeger.
- Westphal, K. R. (1998). Transcendent Reflections on Pragmatic Realism. In K. R. Westphal (Ed.), *Pragmatism, Reason, & Norms: A Realistic Assessment* (17-58). New York: Fordham University Press.
- White, R. T. (1988). *Learning Science*. Oxford, Blackwell.

- White, S. K. (1988). *The Recent Work of Jürgen Habermas*. Cambridge, Cambridge University Press.
- Winograd, T. and F. Flores (1986). *Understanding Computers and Cognition*. Norwood, NJ, Ablex.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford, Blackwell.
- Woods, R. G. and R. S. Barrow (1975). *An Introduction to Philosophy of Education*. London, Methuen.
- Yates, J., M. Bessman, et al. (1988). Are Conceptions of Motion Based on a Naive Theory or on Prototypes? *Cognition* 29 , 251-275.
- Young, I. M. (1997). *Intersecting Voices: Dilemmas of Gender, Political Philosophy and Policy*. Princeton, Princeton University Press.
- Young, R. E. (1988). Critical Teaching and Learning. *Educational Theory* 38 (1), 47-59.
- Young, R. E. (1988). Moral Development, Ego Autonomy, and Questions of Practicality in the Critical Theory of Schooling. *Educational Theory* 38 (4), 391-404.
- Young, R. E. (1990). *A Critical Theory of Education: Habermas and Our Children's Future*. New York, Teachers College Press.
- Young, R. E. (1992). *Critical Theory and Classroom Talk*. Clevedon, Multilingual Matters.